

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions .rnpm and .rnpn. Searches run against the Amino Acid Pending database produce two sets of results, with the extensions .rapm and .rapn.

***Because they contain data that is confidential, the results of Pending database searches should not be left in the case .***

**This Page Blank (uspto)**



## **Protein Sequence Searches - February 2005**

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein, Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 18:52:43 ; Search time 758 Seconds  
(without alignments)  
9949.539 Million cell updates/sec

Title: US-10-679-246-1

Perfect score: 1274

Sequence: 1 ttcttttggttttattgt.....ctcttcgtagtggaagc 1274

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1274	100.0	1274	4	AAC67281 Human Sia
2	1274	100.0	2924	5	Abv28804 Human pro
3	1274	100.0	2924	5	Abv25468 Human pro
4	1274	100.0	2924	5	Abv22970 Human pro
5	1274	100.0	6107	4	Aal05382 Human rep
6	1274	100.0	6107	4	Ab198246 Human tes
7	1272.4	99.9	2829	11	Adm01584 Human cdn
8	1272.4	99.9	2829	13	AdS34362 POSH prot
9	1249	98.0	4090	10	AdF81981 Leukaemia
10	1104	86.7	2128	13	AdS34361 POSH prot
11	1042.8	81.9	2440	11	ADI31262 Human cdn
12	958	75.2	1540	13	AdS34365 POSH prot
13	958	75.2	2048	13	AdS34360 POSH prot
14	957	75.1	2454	13	AdS34359 POSH prot
15	951.4	74.7	1886	13	AdS34359 POSH prot
16	947	74.3	2034	13	AdS34364 POSH prot
17	928.4	72.9	1884	2	Aat64820 Tumour su
18	807.8	63.4	1968	6	Ab199429 Mouse isc
19	761.4	59.8	3128	10	Ad140343 Human pur
20	429	33.7	975	13	ADR25256 Breast ca

21	429	33.7	2502	8	ACC50293	ACC50293 Breast ca
22	427.4	33.5	2240	8	ACC50292	ACC50292 Breast ca
23	427.4	33.5	2240	12	ADP13401	ADP13401 Renal cel
24	427.4	33.5	2240	13	ADR14122	ADR14122 Human NF-
25	427.4	33.5	2240	13	ADP54830	ADP54830 Human NF-
26	425.8	33.4	2198	6	AB199430	AB199430 Mouse isc
27	379	29.7	1535	3	AC98856	AC98856 Human pan
28	354	27.8	4647	4	ABL15923	ABL15923 Drosophil
29	354	27.8	8418	4	ABL15922	ABL15922 Drosophil
30	354	27.8	12029	4	ABL15508	ABL15508 Drosophil
31	293.8	23.1	360	8	ABX44846	ABX44846 Bovine ES
32	268	21.0	466	4	AA187326	AA187326 Human pol
33	191.6	15.0	405	5	ABV38421	ABV38421 Human pro
34	188.2	14.8	426	5	ABV18678	ABV18678 Human pro
35	183.4	14.4	473	5	ABV48461	ABV48461 Human pro
36	183	14.4	277	5	ABV08519	ABV08519 Human pro
37	178	14.0	3349	4	ABL21784	ABL21784 Drosophil
38	161	12.6	257	10	ACA55799	ACA55799 Human sig
39	161	12.6	257	12	ADI55595	ADI55595 Human pol
40	160.2	12.6	411	5	ABV13828	ABV13828 Human pro
41	156.8	12.3	441	5	ABV34940	ABV34940 Human pro
42	145.6	11.4	388	2	AAQ60440	AAQ60440 Human bra
43	139	10.9	417	5	ABV43789	ABV43789 Human pro
44	122.6	9.6	2363	4	ABL21782	ABL21782 Drosophil
45	112	8.8	438	5	ABV04659	ABV04659 Human pro

#### ALIGNMENTS

##### RESULT 1

AAC67281

ID AAC67281 standard; cdna; 1274 BP.

XX

AC AAC67281;

XX

DT 09-APR-2001 (first entry)

XX

DE Human Siah-lalpa coding sequence SEQ ID NO: 1.

XX

KW Human; protein degradation; siah-mediated degradation protein; SMDP;

KW SCF-complex protein; SCP; siah-lalpa; siah-1 interacting protein; SIP;

KW Skpl-associated F-box protein; SAF-1; SAF-2; SAF; cancer; cell division;

KW Skpl-associated destruction-box protein; inflammatory disease; ss.

XX

OS Homo sapiens.

XX

PN WO200077207-A2.

XX

PD 21-DEC-2000.

XX

PF 09-JUN-2000; 2000WO-US015873.

XX

PR 11-JUN-1999; 99US-00330517.

XX

PA (BURN-) BURNHAM INST.

XX

PI Reed JC, Matsuzawa S;

XX

DR WPI; 2001-071273/08.

DR P-PSDB; AAB35157.

XX

PT Siah-Mediated Degradation Protein, useful for drug screening, for therapeutic applications and for functional genomics.

XX

PS Claim 5; Page 95-97; 121pp; English.

XX

CC The present invention provides the protein and coding sequences of several siah-mediated degradation proteins and SCF-complex proteins.

CC These are designated Siah-lalpa, Siah-1 interacting protein (SIP), which

CC encodes two proteins due to alternative splicing (SIP-L and SIP-S), Skpl-

CC associated F-box protein-alpha and beta and -2 (SAF-alpha, SAF-beta

CC and SAF-2) and Skpl-associated destruction-box protein (SAD). The



CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 2924 BP; 871 A; 562 C; 562 G; 925 T; 0 U; 4 Other;

Query Match 100.0%; Score 1274; DB 5; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 1.7e-303;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCTTTAGTGTGTTAGTCCATTTTCTATTTTACCAATTTATTTCTATGTAGTCTAT 60  
DB TTTCTTTAGTGTGTTAGTCCATTTTCTATTTTACCAATTTATTTCTATGTAGTCTAT 539  
QY 61 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTGAAAGAGAGCTTATC 120  
DB CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTGAAAGAGAGCTTATC 599  
QY 121 CAGTGTACAGATCCCTAATAAAGTGCAATTCAGTGTAAATTTTATTTTAAATATCTTTT 180  
DB CAGTGTACAGATCCCTAATAAAGTGCAATTCAGTGTAAATTTTATTTTAAATATCTTTT 659  
QY 181 TTAATCTATTTTCT 240  
DB TTAATCTATTTTCT 719  
QY 241 TATGCGATGTAACATTTATTTAAGTAAGTCATGTTTATTTTATTTTCTCTCTCTCTCT 300  
DB TATGCGATGTAACATTTATTTAAGTAAGTCATGTTTATTTTATTTTCTCTCTCTCTCT 779  
QY 301 CCTTATGTTATTTTATTTTACAAATGAGCGTTCAGATGCTACAGCATTTACCTACCGGTACC 360  
DB CCTTATGTTATTTTATTTTACAAATGAGCGTTCAGATGCTACAGCATTTACCTACCGGTACC 839  
QY 361 TCGAAGTGTCCACCATCCAGAGGGTCTGCGCTGACATGGCACAACTGCATCCAAAT 420  
DB TCGAAGTGTCCACCATCCAGAGGGTCTGCGCTGACATGGCACAACTGCATCCAAAT 899  
QY 421 GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTGTTTGGACTATGTGTACGGCCCATCTTT 480  
DB GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTGTTTGGACTATGTGTACGGCCCATCTTT 959  
QY 481 CAATGTACAGTGGCCATCTGTTTGTAGCAACTGTCGCCAAAGCTCACATGTTGTCCA 540  
DB CAATGTACAGTGGCCATCTGTTTGTAGCAACTGTCGCCAAAGCTCACATGTTGTCCA 1019  
QY 541 ACTTGGCGGGGCCCTTTGGGATCCATTTCCCAACTTGGCTATGGAGAAAGTGGCTAATTC 600  
DB ACTTGGCGGGGCCCTTTGGGATCCATTTCCCAACTTGGCTATGGAGAAAGTGGCTAATTC 1079  
QY 601 GTACTTTTCCCTGTAAATATGCGTCTTCTGATGTAAATTAATCTGTGCCACACAGAA 660  
DB GTACTTTTCCCTGTAAATATGCGTCTTCTGATGTAAATTAATCTGTGCCACACAGAA 1139  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGATGTTAGCCCTTATCTGCTCGGCCCTGGTGTCT 720  
DB AAAGCAGACCATGAAGAGCTCTGTGATGTTAGCCCTTATCTGCTCGGCCCTGGTGTCT 1199  
QY 721 TCTGTAAATGCAAGGCTCTCTGATGCTGTAAATGCCCCCATCTGATGATCAGCATTAAG 780  
DB TCTGTAAATGCAAGGCTCTCTGATGCTGTAAATGCCCCCATCTGATGATCAGCATTAAG 1259  
QY 781 TCCATTACACCTTACAGGAGGATATAGTTTTTCTGCTACAGACATTAATCTTCTCT 840  
DB TCCATTACACCTTACAGGAGGATATAGTTTTTCTGCTACAGACATTAATCTTCTCT 1319  
QY 841 GGTGCTGTGATGGGTGATGATGATGCTCTCTTTTGGCTTTTCACTTCAATGTAGTCTTA 900  
DB GGTGCTGTGATGGGTGATGATGATGCTCTCTTTTGGCTTTTCACTTCAATGTAGTCTTA 1379

QY 901 GAGAAACAGGAAAAATACGATGGTCCACAGAGTCTTTCGCAATCTGTACAGCTGTAGGA 960  
DB GAGAAACAGGAAAAATACGATGGTCCACAGAGTCTTTCGCAATCTGTACAGCTGTAGGA 1439  
QY 961 ACACCCAGCAGCTGMAAATTTTCTTACCGACTTGAGCTTAATGCTCATAGGCGACGA 1020  
DB ACACCCAGCAGCTGMAAATTTTCTTACCGACTTGAGCTTAATGCTCATAGGCGACGA 1499  
QY 1021 TTGACTTTGGGAAGCGACTCTCTCGATCTATTATGAAGGAATTCGAACAGCCATTATGAAT 1080  
DB TTGACTTTGGGAAGCGACTCTCTCGATCTATTATGAAGGAATTCGAACAGCCATTATGAAT 1559  
QY 1081 AGCGACTGTCTAGTCTTTTGACACCGCATGTGCACAGCTTTTTCGAGAAAATGCAATTTA 1140  
DB AGCGACTGTCTAGTCTTTTGACACCGCATGTGCACAGCTTTTTCGAGAAAATGCAATTTA 1619  
QY 1141 GGCATCAATGTAACTATTTTCCATGTTTGAATGGCAATCAAACTTTTCTGCCAGTCT 1200  
DB GGCATCAATGTAACTATTTTCCATGTTTGAATGGCAATCAAACTTTTCTGCCAGTCT 1679  
QY 1201 TTAATACTTCAGTTCACAGAAAATAAGGCACCCCATCTCTGCCAACTTAAACTCTTT 1260  
DB TTAATACTTCAGTTCACAGAAAATAAGGCACCCCATCTCTGCCAACTTAAACTCTTT 1739  
QY 1261 CGGTAGGTGGAAGC 1274  
DB CGGTAGGTGGAAGC 1753  
RESULT 3  
ABV25468  
ID ABV25468 standard; cDNA; 2924 BP.  
XX  
AC ABV25468;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 25459.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PA  
XX Schlegel R, Endege WO, Monahan JE;  
PI  
XX WPI; 2001-662795/76.  
DR  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 5054-5055; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 2924 BP; 871 A; 562 C; 562 G; 925 T; 0 U; 4 Other;  
Query Match 100.0%; Score 1274; DB 5; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 1.7e-303;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTTAGTTGTTATGCTCCATTTTCTATTATTAGCAATTTATTTCTATAGTCTAT 60  
DB 480 TTTCTTTAGTTGTTATGCTCCATTTTCTATTATTAGCAATTTATTTCTATAGTCTAT 539  
QY 61 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120  
DB 540 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 599  
QY 121 CAGGTACAGATCCTAATAAGTGCAATTCAGTGTAAATTTTATTATTTTAAATATCTTTT 180  
DB 600 CAGGTACAGATCCTAATAAGTGCAATTCAGTGTAAATTTTATTATTTTAAATATCTTTT 659  
QY 181 TTAATCCTATTTTCTCTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
DB 660 TTAATCCTATTTTCT 719  
QY 241 TATGGCATGTAACATATTATTAAGTAAGTCAATGTTTATTTTCTCTCTCTCTCTCTCT 300  
DB 720 TATGGCATGTAACATATTATTAAGTAAGTCAATGTTTATTTTCTCTCTCTCTCTCTCT 779  
QY 301 CCTATGTTATTTATTTTCAAGATGAGCGTCCAGCTGTCTACAGCAATTAACCTACCGGTACC 360  
DB 780 CCTATGTTATTTATTTTCAAGATGAGCGTCCAGCTGTCTACAGCAATTAACCTACCGGTACC 839  
QY 361 TCGAAGTGTCCACATCCAGAGGTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 840 TCGAAGTGTCCACATCCAGAGGTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899  
QY 421 GACTTGGCAGTCTTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 900 GACTTGGCAGTCTTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959  
QY 481 CAATGTCCAGAGTGGCCATCTTTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCA 540  
DB 960 CAATGTCCAGAGTGGCCATCTTTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCA 1019  
QY 541 ACTTGGCGGGGCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATTC 600  
DB 1020 ACTTGGCGGGGCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATTC 1079  
QY 601 GTACTTTTCCCTGTAATATGCTCTCTGAGTGTGAATAACTCTGCGCACACACAGAA 660  
DB 1080 GTACTTTTCCCTGTAATATGCTCTCTGAGTGTGAATAACTCTGCGCACACACAGAA 1139  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGTAGGCTTATTCCTGTGCGTCCCTGTGCT 720  
DB 1140 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGTAGGCTTATTCCTGTGCGTCCCTGTGCT 1199  
QY 721 TCCGTAAATGGCAAGCTCTCTGAGTGTGTAATGCCCATCTGTATGATCAGCATCAAG 780  
DB 1200 TCCGTAAATGGCAAGCTCTCTGAGTGTGTAATGCCCATCTGTATGATCAGCATCAAG 1259  
QY 781 TCCATTACAACCTACAGGAGAGGATATAGTTTCTTGTGTCAGACATTAATCTTCTCT 840  
DB 1260 TCCATTACAACCTACAGGAGAGGATATAGTTTCTTGTGTCAGACATTAATCTTCTCT 1319  
QY 841 GGTGCTGTTGACTGGGTGATGATGAGTCCCTGTTTGTGCTTCTACCTCACTGTTAGTCTTA 900

DB 1320 GGTGCTGTTGACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379  
QY 901 GAGAAACAGGAAATACGATGGTCCAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 960  
DB 1380 GAGAAACAGGAAATACGATGGTCCAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 1439  
QY 961 ACAGCGAAGCAAGCTGAAATTTTTCGCTTACCGACTTCAAGCTTAAATGGTCAATAGCGGACGA 1020  
DB 1440 ACAGCGAAGCAAGCTGAAATTTTTCGCTTACCGACTTCAAGCTTAAATGGTCAATAGCGGACGA 1499  
QY 1021 TTGACTTGGGAAGGAGCTCTCTCGATCTATTCATGAAGAAATTCGAAGCAATTAATGAAT 1080  
DB 1500 TTGACTTGGGAAGGAGCTCTCTCGATCTATTCATGAAGAAATTCGAAGCAATTAATGAAT 1559  
QY 1081 AGCGACTCTCTAGTCTTTTGACACAGCAGATTCGACAGCTTTTTCGAGAAATTCGCAATTTA 1140  
DB 1560 AGCGACTCTCTAGTCTTTTGACACAGCAGATTCGACAGCTTTTTCGAGAAATTCGCAATTTA 1619  
QY 1141 GGCATCAATGTAATCTATTTTCCATGTTTGAATGGCAATCAAAATTTTCTGGCCAGTGT 1200  
DB 1620 GGCATCAATGTAATCTATTTTCCATGTTTGAATGGCAATCAAAATTTTCTGGCCAGTGT 1679  
QY 1201 TTAAACTTCTAGTTTTCAGAGAAATTAAGGACCCCATCTGCTGCAACCTAAATCTCTTT 1260  
DB 1680 TTAAACTTCTAGTTTTCAGAGAAATTAAGGACCCCATCTGCTGCAACCTAAATCTCTTT 1739  
QY 1261 CGGTAGGTGGAGC 1274  
DB 1740 CGGTAGGTGGAGC 1753  
RESULT 4  
ABV22970  
ID ABV22970 standard; cDNA; 2924 BP.  
XX AC ABV22970;  
XX DT 13-SEP-2002 (first entry)  
XX DE Human prostate expression marker cDNA 22961.  
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX KW Pharmacogenomic marker; Gene; ss.  
XX OS Homo sapiens.  
XX PN WO200160860-A2.  
XX PD 23-AUG-2001.  
XX PF 20-FEB-2001; 2001WO-US005171.  
XX PR 17-FEB-2000; 2000US-0183319P.  
XX PR 16-MAR-2000; 2000US-0189862P.  
XX PR 25-MAY-2000; 2000US-0207454P.  
XX PR 09-JUN-2000; 2000US-0211314P.  
XX PR 18-JUL-2000; 2000US-0219007P.  
XX PR 13-DEC-2000; 2000US-0255281P.  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer.  
XX Claim 1; Page 4080-4081; 11750pp; English.  
XX The invention relates to an isolated nucleic acid molecule (I) comprising



Qy	1	TTTCTTTAGTGTGTTATGGGCCATTTTCTCTATTTTAGCATTATTATTCTATGAGTGTCTAT	60
Db	3164	TTTCTTTAGTGTGTTATGGGCCATTTTCTCTATTTTAGCATTATTATTCTATGAGTGTCTAT	3223



QY 61 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAAGAGCTTATC 120  
DB 3224 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAAGAGCTTATC 3283  
QY 121 CAGGTACAGATCCTTAATAAGTGACATTCAGTGTAATTTATTTTAAATATCTTTT 180  
DB 3284 CAGGTACAGATCCTTAATAAGTGACATTCAGTGTAATTTATTTTAAATATCTTTT 3343  
QY 181 TTAATCCTATTTTCT 240  
DB 3344 TTAATCCTATTTTCT 3403  
QY 241 TATGGCATGTAACATTTATTAAGTAAGTCAATGTTATTAATTTTCTCTCTCTCTCT 300  
DB 3404 TATGGCATGTAACATTTATTAAGTAAGTCAATGTTATTAATTTTCTCTCTCTCTCT 3463  
QY 301 CCTTATGTTATTTATTTTCAAGATGAGCGTCAAGTCTACAGCATTTACCTACCGGTACC 360  
DB 3464 CCTTATGTTATTTATTTTCAAGATGAGCGTCAAGTCTACAGCATTTACCTACCGGTACC 3523  
QY 361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCTGACTGGCAAACTGTCATCCAAAT 420  
DB 3524 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCTGACTGGCAAACTGTCATCCAAAT 3583  
QY 421 GACTTGGGAGTCTTTTGAAGTCCAGTCTGCTTTGACTGTGTTTACGGCCATCTTT 480  
DB 3584 GACTTGGGAGTCTTTTGAAGTCCAGTCTGCTTTGACTGTGTTTACGGCCATCTTT 3643  
QY 481 CAATGTCAGAGTGGCCATCTTTGTTGTCAGCAACTGCGCCAAAGCTCACATGCTGTCCA 540  
DB 3644 CAATGTCAGAGTGGCCATCTTTGTTGTCAGCAACTGCGCCAAAGCTCACATGCTGTCCA 3703  
QY 541 ACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTC 600  
DB 3704 ACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTC 3763  
QY 601 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATTAATCTCTGCCACACAGAA 660  
DB 3764 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATTAATCTCTGCCACACAGAA 3823  
QY 661 AAAGCAGACCATGAAGAGCTCTGTAGTTTGTAGGCTTATTCCTGTCGTCGCTGCTGTCT 720  
DB 3824 AAAGCAGACCATGAAGAGCTCTGTAGTTTGTAGGCTTATTCCTGTCGTCGCTGCTGTCT 3883  
QY 721 TCCTGTAATGGAAGGCTCTCTGGATGTGTAAATGCCCATCTGATGTCATCAGCATAAG 780  
DB 3884 TCCTGTAATGGAAGGCTCTCTGGATGTGTAAATGCCCATCTGATGTCATCAGCATAAG 3943  
QY 781 TCATTTACAAACCTACAGGAGAGGATATAGTTTTTCTTGCTACACATTAATCTTCCT 840  
DB 3944 TCATTTACAAACCTACAGGAGAGGATATAGTTTTTCTTGCTACACATTAATCTTCCT 4003  
QY 841 GGTGCTGTTGACTGSGTGATGATGAGTCTCTGTTTGGCTTTTCACTTCATGTTAGTCTTA 900  
DB 4004 GGTGCTGTTGACTGSGTGATGATGAGTCTCTGTTTGGCTTTTCACTTCATGTTAGTCTTA 4063  
QY 901 GAGAAACAGGAAATATACGATGTGTACAGCAGTCTCTGCGAATCGTACAGCTGTATAGA 960  
DB 4064 GAGAAACAGGAAATATACGATGTGTACAGCAGTCTCTGCGAATCGTACAGCTGTATAGA 4123  
QY 961 ACACGCAAGCAAGCTGAAATTTTGTCTTACCGACTTTGAGCTTAATGTCATAGGGCAGCA 1020  
DB 4124 ACACGCAAGCAAGCTGAAATTTTGTCTTACCGACTTTGAGCTTAATGTCATAGGGCAGCA 4183  
QY 1021 TTGACTTGGAGCGACTCTCTCGATCTATTCATGAAGGAATTCGAACAGCCATTAAT 1080  
DB 4184 TTGACTTGGAGCGACTCTCTCGATCTATTCATGAAGGAATTCGAACAGCCATTAAT 4243  
QY 1081 AGGACTGTCTAGTCTTTGACACAGCATTTGACAGCTTTTTCAGAAAATGGCAATTTA 1140  
DB 4244 AGGACTGTCTAGTCTTTGACACAGCATTTGACAGCTTTTTCAGAAAATGGCAATTTA 4303  
QY 1141 GGCATCAATGTAATTTTCCATGTGTGTAATGGCAATCAAAATTTTCTGGCCAGTGT 1200

DB 4304 GGCATCAATGTAATTTTCCATGTGTGTAATGGCAATCAAAATTTTCTGGCCAGTGT 4363  
QY 1201 TTAATAACTTTCAGTTCACAGAAAATAAGGACCCCATCTCTGCGCAACCTTAAACTCTTT 1260  
DB 4364 TTAATAACTTTCAGTTCACAGAAAATAAGGACCCCATCTCTGCGCAACCTTAAACTCTTT 4423  
QY 1261 CGGTAGGTGGAAGC 1274  
DB 4424 CGGTAGGTGGAAGC 4437  
RESULT 6  
ABL98246  
ID ABL98246 standard; DNA; 6107 BP.  
XX  
AC ABL98246;  
XX  
DT 21-JUN-2002 (first entry)  
XX  
Human testicular antigen encoding DNA fragment SEQ ID NO: 2898.  
DE  
XX  
Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disease; infection; cytostatic; gene; ds.  
XX  
OS Homo sapiens.  
XX  
WO200155317-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001329.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0209515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0218880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226818P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.

```
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232403P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-023497P.
PR 25-SEP-2000; 2000US-023498P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful
PT for preventing, diagnosing and/or treating testicular cancer.
XX
XX Disclosure; SEQ ID NO 2898; 766pp; English.
XX
XX The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a DNA encoding a
CC protein fragment of the invention
XX
XX Sequence 6107 BP; 1636 A; 1200 C; 1285 G; 1986 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1274; DB 4; Length 6107;
Best Local Similarity 100.0%; Pred. No. 2.1e-303;
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTCTTTAGTTGTTTATGTTCCATTTTCTATTTTATGACATTTATTTCTATAGTCTAT 60
Db 3164 TTCTTTAGTTGTTTATGTTCCATTTTCTATTTTATGACATTTATTTCTATAGTCTAT 3223
QY 61 CCAAGACGATTAAGGAGTCCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 120
Db 3224 CCAAGACGATTAAGGAGTCCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 3283
QY 121 CAGTGTACAGATCCCTAATAAGTGCAATTCAGTGAATTTATTTTATATCTTTT 180
Db 3284 CAGTGTACAGATCCCTAATAAGTGCAATTCAGTGAATTTATTTTATATCTTTT 3343
QY 181 TTAATCCTATTTTCTTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 3344 TTAATCCTATTTTCTTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3403
QY 241 TATGGCATGTAAACATTTATTTATAAGTAAGTCAATGTTTATATTTTCTCTGCCT 300
Db 3404 TATGGCATGTAAACATTTATTTATAAGTAAGTCAATGTTTATATTTTCTCTGCCT 3463
QY 301 CCTTATGTATTTATTTCAAGAAATGAGCGGTACAGCTGCTACAGCATTACCTACCGTACC 360
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful
PT for preventing, diagnosing and/or treating testicular cancer.
XX
XX Disclosure; SEQ ID NO 2898; 766pp; English.
XX
XX The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a DNA encoding a
CC protein fragment of the invention
XX
XX Sequence 6107 BP; 1636 A; 1200 C; 1285 G; 1986 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 1274; DB 4; Length 6107;
Best Local Similarity 100.0%; Pred. No. 2.1e-303;
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTCTTTAGTTGTTTATGTTCCATTTTCTATTTTATGACATTTATTTCTATAGTCTAT 60
Db 3164 TTCTTTAGTTGTTTATGTTCCATTTTCTATTTTATGACATTTATTTCTATAGTCTAT 3223
QY 61 CCAAGACGATTAAGGAGTCCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 120
Db 3224 CCAAGACGATTAAGGAGTCCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 3283
QY 121 CAGTGTACAGATCCCTAATAAGTGCAATTCAGTGAATTTATTTTATATCTTTT 180
Db 3284 CAGTGTACAGATCCCTAATAAGTGCAATTCAGTGAATTTATTTTATATCTTTT 3343
QY 181 TTAATCCTATTTTCTTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 3344 TTAATCCTATTTTCTTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3403
QY 241 TATGGCATGTAAACATTTATTTATAAGTAAGTCAATGTTTATATTTTCTCTGCCT 300
Db 3404 TATGGCATGTAAACATTTATTTATAAGTAAGTCAATGTTTATATTTTCTCTGCCT 3463
QY 301 CCTTATGTATTTATTTCAAGAAATGAGCGGTACAGCTGCTACAGCATTACCTACCGTACC 360
```



1494	Db	 TCCAAAGTGTCCACCATCCAGAGGGTGCCTGCGCCCTGACTGGCACAACTGCATCCCAACAAT	1553
421	Qy	 GACTTGGCGAGTCTTTTTCAGATGTCAGTCTGCTTTTGACTATGTGTGTACCGGCCCATTTCTT	480
1554	Db	 GACTTGGCGAGTCTTTTTCAGATGTCAGTCTGCTTTTGACTATGTGTGTACCGGCCCATTTCTT	1613
481	Qy	 CAATGTCCAGAGTGGCCATCTTTTGTAGCAACTGTGCGCCCAAAAGTCTCATGTGTGTCTCA	540
1614	Db	 CAATGTCCAGAGTGGCCATCTTTTGTAGCAACTGTGCGCCCAAAAGTCTCATGTGTGTCTCA	1673
541	Qy	 ACTTGGCGGGGCCCTTTTGGGATCCATTTCCGAATCTTGGCTATGGAGAAAGTGGCTAATTTCA	600
1674	Db	 ACTTGGCGGGGCCCTTTTGGGATCCATTTCCGAATCTTGGCTATGGAGAAAGTGGCTAATTTCA	1733
601	Qy	 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATAACTCTGCCACACACAGAA	660
1734	Db	 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATAACTCTGCCACACACAGAA	1793
661	Qy	 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGGCCCTTATTTCTGTGCGGTCCCTGTGTCT	720
1794	Db	 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGGCCCTTATTTCTGTGCGGTCCCTGTGTCT	1853
721	Qy	 TCTGTAAATGGCAAGCTCTCTGGATGCTGTAATGCCCATCTGATGATCAATCAGCATAAAG	780
1854	Db	 TCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGATCAATCAGCATAAAG	1913
781	Qy	 TCCATTACAAACCTTACAGGAGAGGATATAGTTTTTCTTGCTACAGACATTAATCTTCTCT	840
1914	Db	 TCCATTACAAACCTTACAGGAGAGGATATAGTTTTTCTTGCTACAGACATTAATCTTCTCT	1973
841	Qy	 GGTGCTGTTGACTGGGTGATGATGCAGTCTGTTTGGCTTTCACATTCATGTTAGTCTTA	900
1974	Db	 GGTGCTGTTGACTGGGTGATGATGCAGTCTGTTTGGCTTTCACATTCATGTTAGTCTTA	2033
901	Qy	 GAGAAACAGGAAAAATACGATGGTCCACGACAGTCTTTCGCAATCGTACAGCTGATAGGA	960
2034	Db	 GAGAAACAGGAAAAATACGATGGTCCACGACAGTCTTTCGCAATCGTACAGCTGATAGGA	2093
961	Qy	 ACAGCAAGCAAGCTGAAAAATTTTGGCTTACCGACTTCGACTTCAGCTAAATGGTCAATGGCAGCA	1020
2094	Db	 ACAGCAAGCAAGCTGAAAAATTTTGGCTTACCGACTTCGACTTCAGCTAAATGGTCAATGGCAGCA	2153
1021	Qy	 TTGACTTTGGGAAGCGACTCCTCGATCTATTTCATCAAGGAATTTGCAACAGGCCATTTATGAAT	1080
2154	Db	 TTGACTTTGGGAAGCGACTCCTCGATCTATTTCATCAAGGAATTTGCAACAGGCCATTTATGAAT	2213
1081	Qy	 AGCGACTGCTCTAGTCTTTTGACACAGCAATTTGCAAGCTTTTTCGAGAAAAATGGCAATTTTA	1140
2214	Db	 AGCGACTGCTCTAGTCTTTTGACACAGCAATTTGCAAGCTTTTTCGAGAAAAATGGCAATTTTA	2273
1141	Qy	 GGCATCAATGTAACTATTTCATCTGTTGAAATGGCAATCAAAATTTTCTGGCCAGTGT	1200
2274	Db	 GGCATCAATGTAACTATTTCATCTGTTGAAATGGCAATCAAAATTTTCTGGCCAGTGT	2333
1201	Qy	 TTAAAACTTTCAGTTTTCAGAAAAATTAAGGCCACCCATCTGTCTGCCAACCTTAAACTCTTTT	1260
2334	Db	 TTAAAACTTTCAGTTTTCAGAAAAATTAAGGCCACCCATCTGTCTGCCAACCTTAAACTCTTTT	2393
1261	Qy	 CGGTAGTGGGAAGC	1274
2394	Db	 CGGTAGTGGGAAGC	2407

RESULT 8	
ADNS34362	
ID	ADNS34362 standard; DNA; 2829 BP.
XX	
XX	
AC	ADNS34362;
XX	
DT	02-DEC-2004 (first entry)
XX	
DE	POSH protein associated DNA #116

ds; gene; cyrostatic; nootropic; neuroprotective; antiparkinsonian; anticonvulsant; antiviral; neuroleptic; central nervous system; POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI; Ubiquitin ligase; anti-viral agent; anti-apoptotic agent; anti-cancer agent; secretory pathway trafficking inhibitor; neurological disorder progression disorder; Alzheimer's disease; Parkinson's disease; Huntington's disease; schizophrenia; Niemann-Pick's disease.

Homo sapiens.

WO2004078130-A2.

16-SEP-2004.

02-MAR-2004; 2004WO-US006308.

03-MAR-2003; 2003US-0451437P.

05-MAR-2003; 2003US-0452284P.

19-MAR-2003; 2003US-0455760P.

20-MAR-2003; 2003US-0456640P.

03-APR-2003; 2003US-0460526P.

04-APR-2003; 2003US-0460782P.

21-APR-2003; 2003US-0464285P.

09-MAY-2003; 2003US-0469462P.

15-MAY-2003; 2003US-0471378P.

20-MAY-2003; 2003US-0472327P.

30-MAY-2003; 2003US-0474706P.

03-JUN-2003; 2003US-0475825P.

17-JUN-2003; 2003US-0479317P.

19-JUN-2003; 2003US-0480215P.

19-JUN-2003; 2003US-0480376P.

08-AUG-2003; 2003US-0493860P.

28-AUG-2003; 2003US-0498634P.

16-SEP-2003; 2003US-0503931P.

10-NOV-2003; 2003WO-US035712.

05-FEB-2004; 2004WO-US003600.

02-MAR-2004; 2004US-0549896P.

(PROT-) PROTEOLOGICS INC.

Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S; Greener T;

WPI; 2004-662346/54.

Isolated, purified or recombinant complex, useful for identifying an antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and POSH-associated protein (POSH-AP).

Disclosure; SEQ ID NO 126; 374pp; English.

The invention relates to an isolated, purified or recombinant complex (1) comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a) or HERPUDI and a ubiquitin ligase (b). Methods using (1), (a) or (b) are useful for identifying an agent that modulates an activity of a POSH polypeptide or POSH-AP, for identifying an antiviral agent, an anti-apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking of a protein through the secretory pathway, an agent that inhibits the progression of a neurological disorder, an agent that modulates a POSH function, an agent that modulates a HERPUDI function. The methods can be used for treating a viral infection, for inhibiting an activity of a POSH-AP in a cell, for treating a POSH-associated disease in a subject. The POSH-associated disease is viral infection, POSH-associated cancer or POSH-associated neurological disorder. The methods are useful for treating or preventing POSH-associated neurological disorder in a subject e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease, schizophrenia, Niemann-Pick's disease. This sequence corresponds to a nucleic acid of the invention.

Sequence 2829 BP; 756 A; 566 C; 553 G; 954 T; 0 U; 0 Other.

Query Match	99.9%;	Score	1272.4;	DB	13;	Length	2829;	
Best Local Similarity	99.9%;	Pred. No.	4e-303;					
Matches 1273;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
1	TTTCTTTAGTCTGTTATAGTGCCATTTCTATTTTATAGCATTTATTTATTTCTATGTAAGTCTAT	60						
1134	TTTCTTTAGTCTGTTATAGTGCCATTTCTATTTTATAGCATTTATTTATTTCTATGTAAGTCTAT	1193						
61	CCAAAGACGATTAAGGGAGTTCACATGTTTTTCCGGAACTATTTTGAAAGAGAGCTTATC	120						
1194	CCAAAGACGATTAAGGGAGTTCACATGTTTTTCCGGAACTATTTTGAAAGAGAGCTTATC	1253						
121	CAGTGTACAGATCCCTAAATAAGTGCACATTCAGTGTAAATTTTATTTTAAATATCTTTT	180						
1254	CAGTGTACAGATCCCTAAATAAGTGCACATTCAGTGTAAATTTTATTTTAAATATCTTTT	1313						
181	TTAAATCCATTTTCTCTCTCTTTTGTCTCAGTAAATTTTGTATGAACTTTTAAAGAGACT	240						
1314	TTAAATCCATTTTCTCTCTCTTTTGTCTCAGTAAATTTTGTATGAACTTTTAAAGAGACT	1373						
241	TATGGCATGTAAACATTTATTTATAAGTAAGTTCATGGTTATAATTTATTTTCTCTCGCT	300						
1374	TATGGCATGTAAACATTTATTTATAAGTAAGTTCATGGTTATAATTTATTTTCTCTCGCT	1433						
301	CCTTATGTATTTATTTTCAGAAATGAGCGTGACCTGTACAGCATTTACTTACCGGTACC	360						
1434	CCTTATGTATTTATTTTCAGAAATGAGCGTGACCTGTACAGCATTTACTTACCGGTACC	1493						
361	TCGAGTGTCCACCATCCGAGGGTGCTCCCTGTGCTGACCTGACCACTGCAATCCAAAT	420						
1494	TCGAAAGTGTCCACCATCCGAGGGTGCTCCCTGTGCTGACCTGACCACTGCAATCCAAAT	1553						
421	GACTTGGCGAGTCTTTTGTGAGTGTCCAGTCTGTCTTGTACTATGTGTACCGCCATCTT	480						
1554	GACTTGGCGAGTCTTTTGTGAGTGTCCAGTCTGTCTTGTACTATGTGTACCGCCATCTT	1613						
481	CAATGTACAGTGGCCATCTTTTGTGAGCAACTGTGTGCCCAAGCTCACATGTTGTCCA	540						
1614	CAATGTACAGTGGCCATCTTTTGTGAGCAACTGTGTGCCCAAGCTCACATGTTGTCCA	1673						
541	ACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAAGTGGCTAATTTCA	600						
1674	ACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAAGTGGCTAATTTCA	1733						
601	GTACTTTTTCCCCTGTAAATATGCGGTCTTCTGGATGTGAAATAAATCTGTGCCACACACAGAA	660						
1734	GTACTTTTTCCCCTGTAAATATGCGGTCTTCTGGATGTGAAATAAATCTGTGCCACACACAGAA	1793						
661	AAAGCAGACCATGAAGAGCTCTGTGAGTTTATGGCCCTTATTCCTGTCCGTGCCGTGGTCT	720						
1794	AAAGCAGACCATGAAGAGCTCTGTGAGTTTATGGCCCTTATTCCTGTCCGTGCCGTGGTCT	1853						
721	TCCTGTAAATGGCAAGGCTCTCTGTGATGCTGTAAATGCCCCATCTGTATGTCATCAGCATTAAG	780						
1854	TCCTGTAAATGGCAAGGCTCTCTGTGATGCTGTAAATGCCCCATCTGTATGTCATCAGCATTAAG	1913						
781	TCCATTTACAACTTACGGAGAGGATATAGTTTTTCTTGTGCTACAGACATTTAATCTTCCT	840						
1914	TCCATTTACAACTTACGGAGAGGATATAGTTTTTCTTGTGCTACAGACATTTAATCTTCCT	1973						
841	GGTGTCTGTGATGGGTGATGATGCAAGTCTGTGTTTGGCTTTTCACTTCATGTTAGTCTTA	900						
1974	GGTGTCTGTGATGGGTGATGATGCAAGTCTGTGTTTGGCTTTTCACTTCATGTTAGTCTTA	2033						
901	GAGAAACAGGAAAAATACGATGGTCCACAGAGTCTTCTTGGCAATTCGTAAGCTGATAGGA	960						
2034	GAGAAACAGGAAAAATACGATGGTCCACAGAGTCTTCTTGGCAATTCGTAAGCTGATAGGA	2093						
961	ACAGCAAGCAGCTGAAAAATTTTGTCTTACCGACTTGCAGCTTAAATGTCATAGCGCAGCA	1020						
2094	ACAGCAAGCAGCTGAAAAATTTTGTCTTACCGACTTGCAGCTTAAATGTCATAGCGCAGCA	2153						
1021	TTGACTTGGGAAGCGACTCTCGATCTTATTCATGAAGAAATGTCACAGCCATTTATGAAT	1080						

Db	2154	TTGACTTTGGGAAGCGACTCTCGATCTATTTCATGAGGAAATTGCAACAGGCATTATGAAT	2213
Qy	1081	AGCGACTGCTTAGTCTTTTGACACCGACCATTCGACAGCTTTTTCGAGAAAATGGCAATTTTA	1140
Db	2214	AGCGACTGCTTAGTCTTTTGACACCGACCATTCGACAGCTTTTTCGAGAAAATGGCAATTTTA	2273
Qy	1141	GGCATCAATGTAACTATTTCATGTGTGTAATGGCAATCAAAACATTTTCTGGCCAGTGT	1200
Db	2274	GGCATCAATGTAACTATTTCATGTGTGTAATGGCAATCAAAACATTTTCTGGCCAGTGT	2333
Qy	1201	TTAAACTTCAGTTTTCACAGAAAATAAGGCACCCATCTGTCTGCCAACCTTAAACTCTTTT	1260
Db	2334	TTAAACTTCAGTTTTCACAGAAAATAAGGCACCCATCTGTCTGCCAACCTTAAACTCTTTT	2393
Qy	1261	CGGTAGTGGAAGC	1274
Db	2394	CGGTAGTGGAAGC	2407
RESULT 9			
ADF81981/c			
ID	ADF81981 standard; DNA; 4090 BP.		
XX	AC		
XX	AC		
XX	AC		
DT	26-FEB-2004 (first entry)		
XX	Leukaemia-related DNA sequence #2537.		
DE	Leukaemia-related DNA sequence #2537.		
XX	Cytostatic; Gene therapy; leukaemia; ss.		
KW	Unidentified.		
OS	WO2003039443-A2.		
PN	15-MAY-2003.		
XX	04-NOV-2002; 2002WO-EP012303.		
PF	05-NOV-2001; 2001EP-00126244.		
XX	30-APR-2002; 2002EP-00009758.		
PR	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.		
XX	(UYLU-) UNIV LUDWIG MAXIMILIANS.		
PA	(HAFE/) HAFERLACH T.		
PA	(SCHO/) SCHOCH C.		
PA	(KERN/) KERN W.		
XX	Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;		
PI	Eils R, Brors B, Mergenthaler S;		
XX	WPI; 2003-505037/47.		
DR	Determining the subtype of leukemia cells and whether a patient sample		
XX	contains leukemia cells or other cells, useful for treating leukemia,		
PT	comprises determining the expression profile of a group of markers in a		
PT	patient sample.		
XX	Disclosure; SEQ ID NO 2537; 2938pp; English.		
PS	The present invention relates to a method (M1) for determining the		
XX	subtype of leukaemia cells and whether a patient sample contains		
CC	leukaemia cells. The method comprises determining the expression profile		
CC	of a group of markers in a patient sample. The method is useful for		
CC	determining the presence of leukaemia cells, its types or subtypes, and		
CC	for the preparation of a medicament for treating leukaemia.		
XX	Sequence 4090 BP; 1233 A; 840 C; 768 G; 1189 T; 0 U; 60 Other;		
SQ	Query Match 98.0%; Score 1249; DB 10; Length 4090;		
XX	Best Local Similarity 99.6%; Pred. No. 2.7e-297;		
XX	Matches 1271; Conservative 0; Mismatches 3; Indels 2; Gaps 2;		

Qy	1	TTTCTTTAGTTGTTTATGGTCCATTTTCTATTTTATTTAGCATTTTATTTCTATGTAGTCTAT	60	Qy	1079	ATAGCGAGTCTGTAGTCTTTTGACACACAGCATTTGCACAGCTTTTTCGAGAAATGGCAATT	1138		
Ds	3611	TTTCTTTAGTTGTTTATGGTCCATTTTCTATTTTATTTAGCATTTTATTTCTATGTAGTCTAT	3552	Ds	2531	ATAGCGAGTCTGTAGTCTTTTGACACACAGCATTTGCACAGCTTTTTCGAGAAATGGCAATT	2472		
Qy	61	CCAAAGACGATTAAGGGAGTTCACATGTTTTCGGAAACATTTTGAAAGAGAGCTTATC	120	Qy	1139	TAGGCATCAATGTAACATTTTCCATGTGTTGAAATGGCAATCAACATTTTCTGGCCAGT	1198		
Ds	3551	CCAAAGACGATTAAGGGAGTTCACATGTTTTCGGAAACATTTTGAAAGAGAGCTTATC	3492	Ds	2471	TAGGCATCAATGTAACATTTTCCATGTGTTGAAATGGCAATCAACATTTTCTGGCCAGT	2412		
Qy	121	CAGTGTACAGATCCCTAATAAGTGCACATTCAGTGTAAATTTTATTTTATATCTTTT	180	Qy	1199	GTTTAAACTTCAGTTTTCACAGAAATAAGGCCACCCATCTGTCTGCCAACCTAAAACTCT	1258		
Ds	3491	CAGTGTACAGATCCCTAATAAGTGCACATTCAGTGTAAATTTTATTTTATATCTTTT	3432	Ds	2411	GTTTAAACTTCAGTTTTCACAGAAATAAGGCCACCCATCTGTCTGCCAACCTAAAACTCT	2352		
Qy	181	TTAATCCTATTTTCTCTCTTTTCTCAGTAAATTTTGTATGAACCTTTTAAAGGACT	240	Qy	1259	TTCGGTAGGTGGAAGC	1274		
Ds	3431	TTAATCCTATTTTCTCTCTTTTCTCAGTAAATTTTGTATGAACCTTTTAAAGGACT	3372	Ds	2351	TTCGGTAGGTGGAAGC	2336		
Qy	241	TATGGCATGTAACATTTATTTAAGTAAGTCAATGTTTATTAATTTTCTCTCGCT	300	RESULT 10					
Ds	3371	TATGGCATGTAACATTTATTTAAGTAAGTCAATGTTTATTAATTTTCTCTCGCT	3312	ADS34361/c					
Qy	301	CCTTATGTTATTTTTCAGAAATGAGCCGTACAGCTGTCTACAGCAATACCTACCGGTACC	360	ID ADS34361 standard; DNA; 2128 BP.					
Ds	3311	CCTTATGTTATTTTTCAGAAATGAGCCGTACAGCTGTCTACAGCAATACCTACCGGTACC	3252	XX	ADS34361;				
Qy	361	TCGAAGTGTCCACATCCACAGAGGGTCCCTGCCCTGACTGGCACAACTGCATCCCAAT	420	XX	02-DEC-2004 (first entry)				
Ds	3251	TCGAAGTGTCCACATCCACAGAGGGTCCCTGCCCTGACTGGCACAACTGCATCCCAAT	3192	XX	POSH protein associated DNA #115.				
Qy	421	GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCCATCTT	480	ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;					
Ds	3191	GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCCATCTT	3132	anticonvulsant; antiviral; neuroleptic; central nervous system;					
Qy	481	CAATGTACAGAGTGGCCATCTTGTGTAGCAACTGTGCGCCAAAGTCAATGTGTCCA	540	POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;					
Ds	3131	CAATGTACAGAGTGGCCATCTTGTGTAGCAACTGTGCGCCAAAGTCAATGTGTCCA	3072	Ubiquitin ligase; antiviral agent; anti-apoptotic agent;					
Qy	541	ACTTGGCGGGCCCTTTGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAATCCA	600	anti-cancer agent; secretory pathway trafficking inhibitor;					
Ds	3071	ACTTGGCGGGCCCTTTGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAATCCA	3012	neurological disorder progression disorder; Alzheimer's disease;					
Qy	601	GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATAACTCTGCCACACACAGAA	660	Parkinson's disease; Huntington's disease; schizophrenia;					
Ds	3011	GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATAACTCTGNCACACACAGAA	2952	Niemann-Pick's disease.					
Qy	661	AAA-GCAGACCATGAAGA-GCTCTGTGAGTTTATGGCTTATTCCTGCGTCCCTGGTG	718	Homo sapiens.					
Ds	2951	AAANGCAGACCATGAANANGCTCTGTGAGTTTATGGCTTATTCCTGCGTCCCTGGTG	2892	WO2004078130-A2.					
Qy	719	CTTCTCTGTAATGGCAAGCTCTCTGGATGCTGTATGCCCCCATCTGATGCATCAGCATA	778	16-SEP-2004.					
Ds	2891	CTTCTCTGTAATGGCAAGCTCTCTGGATGCTGTATGCCCCCATCTGATGCATCAGCATA	2832	02-MAR-2004; 2004WO-US0006308.					
Qy	779	AGTCCATTACAAACCTTACAGGAGAGGATATAGTTTTTCTGTACAGCAATTAATCTTC	838	03-MAR-2003; 2003US-0451437P.					
Ds	2831	AGTCCATTACAAACCTTACAGGAGAGGATATAGTTTTTCTGTACAGCAATTAATCTTC	2772	05-MAR-2003; 2003US-0452284P.					
Qy	839	CTGTGCTCTTGA CTGGGTGATGATGCAGTCTGTTTTTGGCTTTTCACTTCATGTTAGTCT	898	19-MAR-2003; 2003US-0455760P.					
Ds	2771	CTGGTCTCTTGA CTGGGTGATGATGCAGTCTGTTTTTGGCTTTTCACTTCATGTTAGTCT	2712	20-MAR-2003; 2003US-0456640P.					
Qy	899	TAGAGAAACAGGAAATATACGATGGTCCACAGCAGTTCTTCGCAATCGTACAGCTGATAG	958	03-APR-2003; 2003US-0460526P.					
Ds	2711	TAGAGAAACAGGAAATATACGATGGTCCACAGCAGTTCTTCGCAATCGTACAGCTGATAG	2652	04-APR-2003; 2003US-0460792P.					
Qy	959	GAACACGCAAGCAAGCTGAAATTTTGGCTTACCGACTTGAGCTAAATGGTCATAGGCGAC	1018	21-APR-2003; 2003US-0464285P.					
Ds	2651	GAACACGCAAGCAAGCTGAAATTTTGGCTTACCGACTTGAGCTAAATGGTCATAGGCGAC	2592	09-MAY-2003; 2003US-0469462P.					
Qy	1019	GATTGACTTGGGAGCGACTCTCGATCTATTTCATGAAGGAATTCGAACAGCCATTATGA	1078	15-MAY-2003; 2003US-0471378P.					
Ds	2591	GATTGACTTGGGAGCGACTCTCTCGATCTATTTCATGAAGGAATTCGAACAGCCATTATGA	2532	20-MAY-2003; 2003US-0472327P.					
				30-MAY-2003; 2003US-0474706P.					
				03-JUN-2003; 2003US-0475825P.					
				17-JUN-2003; 2003US-0479311P.					
				19-JUN-2003; 2003US-0480215P.					
				19-JUN-2003; 2003US-0480376P.					
				08-AUG-2003; 2003US-0493860P.					
				28-AUG-2003; 2003US-0498634P.					
				16-SEP-2003; 2003US-0503931P.					
				10-NOV-2003; 2003WO-US035712.					
				05-FEB-2004; 2004WO-US003600.					
				02-MAR-2004; 2004US-0549896P.					
				(PROT-) PROTEOLOGICS INC.					
				Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;					
				Greener T;					
				WPI; 2004-662346/64.					





CC The invention relates to a composition comprising a plurality of cDNAs  
CC for detecting the altered expression of genes in an immunological  
CC response. The invention also relates to a method of diagnosing or  
CC monitoring the treatment of an immunopathological condition in a sample,  
CC comprising obtaining nucleic acids from a sample, contacting the nucleic  
CC acids of the sample with an array comprising the plurality of cDNAs under  
CC conditions to form one or more hybridisation complexes, detecting the  
CC hybridisation complexes and comparing the levels of the detected  
CC hybridisation complexes with the level of hybridisation complexes  
CC detected in a non-diseased sample, where an altered level of the detected  
CC hybridisation complexes correlates with the presence of an  
CC immunopathological condition. Also disclosed are an expression profile  
CC comprising a microarray and a plurality of detectable complexes and a  
CC method for identifying a plurality of polynucleotide probes. The cDNAs  
CC are useful as hybridisable array elements in a microarray for monitoring  
CC the expression of target polynucleotides. The microarray can be used in  
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,  
CC ulcerative colitis, hyperosinophilia, irritable bowel syndrome,  
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in  
CC identifying agents for the treatment of the diseases. The microarray may  
CC also be used in drug discovery and development, toxicological and  
CC carcinogenicity studies, forensics or pharmacogenomics. The composition  
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or  
CC genomic fragments. This sequence represents a human cDNA of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification but was obtained in electronic format directly  
CC from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX SQ Sequence 2440 BP; 685 A; 463 C; 472 G; 817 T; 0 U; 3 Other;

Query Match 81.9%; Score 1042.8; DB 11; Length 2440;  
Best Local Similarity 99.8%; Pred. No. 1.5e-246;  
Matches 1044; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	229	TTTTAAAGGACCTTATGGCATGTAACATTATTTATTAAGTAAGTCAATGTTATTAATT	288
DB	462	TTTTGGAAGGACCTTATGGCATGTAACATTATTTATTAAGTAAGTCAATGTTATTAATT	521
QY	289	TTTCTCTGCTCCTTATGTTATTTATTTTCAAGTAAGTCAATGTTATTAATT	348
DB	522	TTTCTCTGCTCCTTATGTTATTTATTTTCAAGTAAGTCAATGTTATTAATT	581
QY	349	CTTACCGGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTCCCTGAGTGGCAAACT	408
DB	582	CTTACCGGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTCCCTGAGTGGCAAACT	641
QY	409	GCATCCCAACAAATGACTTGGCGAGTCTTTTGGAGTCCAGTCTGCTTGAATGTTA	468
DB	642	GCATCCCAACAAATGACTTGGCGAGTCTTTTGGAGTCCAGTCTGCTTGAATGTTA	701
QY	469	CCGCCCAATCTTCAATGTCAGAGTGGCCATCTTTTGGAGTCCAGTCTGCTTGAATGTTA	528
DB	702	CCGCCCAATCTTCAATGTCAGAGTGGCCATCTTTTGGAGTCCAGTCTGCTTGAATGTTA	761
QY	529	ACATGTTGTCCAACTTCCCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAA	588
DB	762	ACATGTTGTCCAACTTCCCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAA	821
QY	589	GTGGCTAATTCAGTACTTTTCCCTCTTAATATGCGTCTTCTGGATGTGAATACTCTG	648
DB	822	GTGGCTAATTCAGTACTTTTCCCTCTTAATATGCGTCTTCTGGATGTGAATACTCTG	881
QY	649	CCACACAGAAAAGCAGACCATGAGAGCTCTGTGAGTCTTGGGCTTATTCCTGTCG	708
DB	882	CCACACAGAAAAGCAGACCATGAGAGCTCTGTGAGTCTTGGGCTTATTCCTGTCG	941
QY	709	TGCCCTGTGCTCTCTGTTAAATGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATG	768
DB	942	TGCCCTGTGCTCTCTGTTAAATGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATG	1001
QY	769	CATCAGCATAGTCCATTTACAACTTACAGGAGGATATAGTTTTCTTGTCTACAGAC	828
DB	1002	CATCAGCATAGTCCATTTACAACTTACAGGAGGATATAGTTTTCTTGTCTACAGAC	1061

QY	829	ATTAAATCTTCTGCTGCTGTTGACTGGGTGATGATGAGTCTGCTGTTTGGCTTTCACCTC	888
DB	1062	ATTAAATCTTCTGCTGCTGTTGACTGGGTGATGATGAGTCTGCTGTTTGGCTTTCACCTC	1121
QY	889	ATGTTAGTCTTTAGAGAAACAGGAAAATAGATGGTCCACGAGTCTTTCGCAATCGTA	948
DB	1122	ATGTTAGTCTTTAGAGAAACAGGAAAATAGATGGTCCACGAGTCTTTCGCAATCGTA	1181
QY	949	CAGCTGATAGAAACACCAAGCAAGCTGAAATTTTCTTACCGACTTGAGCTAAATGGT	1008
DB	1182	CAGCTGATAGAAACACCAAGCAAGCTGAAATTTTCTTACCGACTTGAGCTAAATGGT	1241
QY	1009	CATAGGGACCAATTTGACTTTGGGAAGCGACTCTCGATCTATTATGAGGAATTCGAACA	1068
DB	1242	CATAGGGACCAATTTGACTTTGGGAAGCGACTCTCGATCTATTATGAGGAATTCGAACA	1301
QY	1069	GCCATTATGAAATAGGACTGCTAGTCTTTTACACACGAGTTCACAGCTTTTTCGAGAA	1128
DB	1302	GCCATTATGAAATAGGACTGCTAGTCTTTTACACACGAGTTCACAGCTTTTTCGAGAA	1361
QY	1129	AATGCAATTTAGGATCAATGTAATTTTCCATGTTTGAATGGCAATCAACATTT	1188
DB	1362	AATGCAATTTAGGATCAATGTAATTTTCCATGTTTGAATGGCAATCAACATTT	1421
QY	1189	TCTGCGCAGTGTTTAAACCTTCAAGTTTACAGAAATAAGGACCCATCTCTGCTGCCAAC	1248
DB	1422	TCTGCGCAGTGTTTAAACCTTCAAGTTTACAGAAATAAGGACCCATCTCTGCTGCCAAC	1481
QY	1249	CTAAACCTTTTCGGTAGGTGGAAGC	1274
DB	1482	CTAAACCTTTTCGGTAGGTGGAAGC	1507

RESULT 12  
ADS34365  
ID ADS34365 standard; DNA; 1540 BP.  
XX AC ADS34365;  
XX DT 02-DEC-2004 (first entry)  
XX DE POSH protein associated DNA #119.  
XX ds; gene; cytostatic; neurotropic; neuroprotective; antiparkinsonian;  
XX anticonvulsant; antiviral; neuroleptic; central nervous system;  
XX POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;  
XX Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
XX anti-cancer agent; secretory pathway trafficking inhibitor;  
XX neurological disorder progression disorder; Alzheimer's disease;  
XX Parkinson's disease; Huntington's disease; schizophrenia;  
XX Niemann-Pick's disease.  
XX Homo sapiens.  
XX WO2004078130-A2.  
XX PD 16-SEP-2004.  
XX PF 02-MAR-2004; 2004WO-US006308.  
XX PR 03-MAR-2003; 2003US-0451437P.  
XX PR 05-MAR-2003; 2003US-0452284P.  
XX PR 19-MAR-2003; 2003US-0455760P.  
XX PR 20-MAR-2003; 2003US-0456640P.  
XX PR 03-APR-2003; 2003US-0460526P.  
XX PR 04-APR-2003; 2003US-0460792P.  
XX PR 21-APR-2003; 2003US-0464285P.  
XX PR 09-MAY-2003; 2003US-0469462P.  
XX PR 15-MAY-2003; 2003US-0471378P.  
XX PR 20-MAY-2003; 2003US-0472327P.  
XX PR 30-MAY-2003; 2003US-0474706P.  
XX PR 03-JUN-2003; 2003US-0475825P.





[illegible]

Qy	558	GGGATCATTTCGCAACTTTGGCTATGGAGAAAGTGGCTAAATCTAGTACTTTTCCCTCTGTA	617
Db	696	GGGATCATTTCGCAACTTTGGCTATGGAGAAAGTGGCTAAATCTAGTACTTTTCCCTCTGTA	755
Qy	618	ATATGCGCTCTTCGGATGTGAAATACTCTGCCACACACAGAAAAGCAGACCATGAAGA	677
Db	756	ATATGCGCTCTTCGGATGTGAAATACTCTGCCACACACAGAAAAGCAGACCATGAAGA	815
Qy	678	GCTCTGTGAGTTTAGGCTTTAATCTCTGTCGTCGCCCTGGTCTTCCCTGTAATATGGCAAG	737
Db	816	GCTCTGTGAGTTTAGGCTTTAATCTCTGTCGTCGCCCTGGTCTTCCCTGTAATATGGCAAG	875
Qy	738	CTCTCTGGATGCTGTAATGCCCCATCTGATGTCATCAGCATAGTCCATTAACAACCTTACA	797
Db	876	CTCTCTGGATGCTGTAATGCCCCATCTGATGTCATCAGCATAGTCCATTAACAACCTTACA	935
Qy	798	GGGAGAGGATATAGTTTTTCTCTGCTACAGACATTAATCTTCTCTGGTGTCTGTGACTGGGT	857
Db	936	GGGAGAGGATATAGTTTTTCTCTGCTACAGACATTAATCTTCTCTGGTGTCTGTGACTGGGT	995
Qy	858	GATGATGCACTCTGTTTTTGGCTTTCACTTCATGTTAGTCTTAGAGAAAACAGGAAAAATA	917
Db	996	GATGATGCACTCTGTTTTTGGCTTTCACTTCATGTTAGTCTTAGAGAAAACAGGAAAAATA	1055
Qy	918	CGATGGTCACGACGATCTTCTCCGAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGA	977
Db	1056	CGATGGTCACGACGATCTTCTCCGAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGA	1115
Qy	978	AAATTTTGTCTTACCGACTTGAGCTAAATGGTTCATAGCCGACGATTCAGTTGGGAAGCGAC	1037
Db	1116	AAATTTTGTCTTACCGACTTGAGCTAAATGGTTCATAGCCGACGATTCAGTTGGGAAGCGAC	1175
Qy	1038	TCCTCGATCTATTATGAAGGAATTGCAACAGGCCATATGAAATAGGCACTGTCTAGTCTT	1097
Db	1176	TCCTCGATCTATTATGAAGGAATTGCAACAGGCCATATGAAATAGGCACTGTCTAGTCTT	1235
Qy	1098	TGACACGACATTCACAGCTTTTTCAGAAAAATGCAATTTTAGGGCATCAATGTAACAT	1157
Db	1236	TGACACGACATTCACAGCTTTTTCAGAAAAATGCAATTTTAGGGCATCAATGTAACAT	1295
Qy	1158	TTCCATGTGTGAAATGGCAATCAAAATTTTCTGCGCAGTGTGTTAAAACTTCAGTTTCA	1217
Db	1296	TTCCATGTGTGAAATGGCAATCAAAATTTTCTGCGCAGTGTGTTAAAACTTCAGTTTCA	1355
Qy	1218	CAGAAAATATAGGCACCCATCTGTCTGCCAACCTTAAACCTTTTCGGTAGTGGAAAGC	1274
Db	1356	CAGAAAATATAGGCACCCATCTGTCTGCCAACCTTAAACCTTTTCGGTAGTGGAAAGC	1412
RESULT 15			
ADS34363			
ID	ADS34363 standard; DNA; 1886 BP.		
XX	AC		
XX	ADS34363;		
DT	02-DEC-2004 (first entry)		
XX	POSH protein associated DNA #117.		
DE	ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;		
XX	anticonvulsant; antiviral; neuroleptic; central nervous system;		
KW	POSH polypeptide; POSH-associated protein; POSH-AP; HEPUD1;		
KW	Ubiquitin ligase; antiviral agent; anti-apoptotic agent;		
KW	anti-cancer agent; secretory pathway trafficking inhibitor;		
KW	neurological disorder progression disorder; Alzheimer's disease;		
KW	Parkinson's disease; Huntington's disease; schizophrenia;		
XX	Niemann-Pick's disease.		
OS	Homo sapiens.		
XX	Wo2004078130-A2.		
PN	XX		

```
XX 02-MAR-2004; 2004WO-US006308.
XX PF
XX 03-MAR-2003; 2003US-0451437P.
XX PR
XX 05-MAR-2003; 2003US-0452284P.
XX PR
XX 19-MAR-2003; 2003US-0455760P.
XX PR
XX 20-MAR-2003; 2003US-0456640P.
XX PR
XX 03-APR-2003; 2003US-0460526P.
XX PR
XX 04-APR-2003; 2003US-0460792P.
XX PR
XX 21-APR-2003; 2003US-0464285P.
XX PR
XX 09-MAY-2003; 2003US-0469462P.
XX PR
XX 15-MAY-2003; 2003US-0471378P.
XX PR
XX 20-MAY-2003; 2003US-0472327P.
XX PR
XX 30-MAY-2003; 2003US-0474706P.
XX PR
XX 03-JUN-2003; 2003US-0475825P.
XX PR
XX 17-JUN-2003; 2003US-0479317P.
XX PR
XX 19-JUN-2003; 2003US-0480215P.
XX PR
XX 19-JUN-2003; 2003US-0480376P.
XX PR
XX 08-AUG-2003; 2003US-0493860P.
XX PR
XX 28-AUG-2003; 2003US-0498634P.
XX PR
XX 16-SEP-2003; 2003US-0503931P.
XX PR
XX 10-NOV-2003; 2003WO-US035712.
XX PR
XX 05-FEB-2004; 2004WO-US003600.
XX PR
XX 02-MAR-2004; 2004US-0549896P.
XX PR
XX (PROT-) PROTEOLOGICS INC.
XX PI
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
XX PI Greener T;
XX WPI; 2004-662346/64.
XX DR
XX Isolated, purified or recombinant complex, useful for identifying an
XX PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
XX PT POSH-associated protein (POSH-AP).
XX PS
XX Disclosure; SEQ ID NO 127; 374bp; English.
XX CC
XX The invention relates to an isolated, purified or recombinant complex (I)
XX CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
XX CC or HERPUL1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are
XX CC useful for identifying an agent that modulates an activity of a POSH
XX CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
XX CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
XX CC of a protein through the secretory pathway, an agent that inhibits the
XX CC progression of a neurological disorder, an agent that modulates a POSH
XX CC function, an agent that modulates a HERPUL1 function. The methods can be
XX CC used for treating a viral infection, for inhibiting an activity of a POSH
XX CC -AP in a cell, for treating a POSH-associated disease in a subject. The
XX CC POSH-associated disease is viral infection, POSH-associated cancer or
XX CC POSH-associated neurological disorder. The methods are useful for
XX CC treating or preventing POSH-associated neurological disorder in a subject
XX CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
XX CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
XX CC nucleic acid of the invention.
XX CC
XX SQ
XX Sequence 1886 BP; 527 A; 366 C; 373 G; 620 T; 0 U; 0 Other;
XX
XX Query Match 74.7%; Score 951.4; DB 13; Length 1886;
XX Best Local Similarity 99.9%; Pred. No. 4.5e-224;
XX Matches 952; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 322 ATGAGCGCTGACAGCTGTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCAG 381
XX DB 1 ATGAGCGCTGACAGCTGTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCAG 60
XX
XX 382 AGGCTGCTGCCCTGACTGGCACAACATGCTATCCAAATGACTTGGCGAGTCTTTTGAG 441
XX DB 61 AGGCTGCTGCCCTGACTGGCACAACATGCTATCCAAATGACTTGGCGAGTCTTTTGAG 120
XX
XX 442 TGTCCAGTCTGCTTGTACTATGTGTACCGCCCAATCTTCAATGTCAGAGTGGCCATCTT 501
XX DB 121 TGTCCAGTCTGCTTGTACTATGTGTACCGCCCAATCTTCAATGTCAGAGTGGCCATCTT 180
```

```
QY 502 GTTTGTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCAACTTGCAGGCGCCCTTTGGGA 561
DB 181 GTTTGTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCAACTTGCAGGCGCCCTTTGGGA 240
QY 562 TCCATTCGCAACTTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTCTAAATAT 621
DB 241 TCCATTCGCAACTTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTCTAAATAT 300
QY 622 GCCTCTTCTGGAGTGTAAATACTCTGCCACACACAGAAAGAGCAGACCATGAAGAGCTC 681
DB 301 GCCTCTTCTGGAGTGTAAATACTCTGCCACACACAGAAAGAGCAGACCATGAAGAGCTC 360
QY 682 TGTGAGTTTAGGCTTATTCCTGTCCTGCGCTTCTTCTGCTTCTTCTGCTTCTTCTGCTTCT 741
DB 361 TGTGAGTTTAGGCTTATTCCTGTCCTGCGCTTCTTCTGCTTCTTCTGCTTCTTCTGCTTCT 420
QY 742 CTGATGCTGTAATGCGCCCATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 801
DB 421 CTGATGCTGTAATGCGCCCATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 480
QY 802 GAGGATAGTCTTTTCTTGTCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 861
DB 481 GAGGATAGTCTTTTCTTGTCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 862 ATGCACTCTCTTTTGGCTTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921
DB 541 ATGCACTCTCTTTTGGCTTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 922 GGTACACGAGCTTCTTTCGCAATGCTGATGAGGAAACACGCAAGCTGAAAT 981
DB 601 GGTACACGAGCTTCTTTCGCAATGCTGATGAGGAAACACGCAAGCTGAAAT 660
QY 982 TTTGCTTACGACTTGTAGCTAAATGCTCATAGGCGACGATTTGACTTGGGAGCGACTCTCT 1041
DB 661 TTTGCTTACGACTTGTAGCTAAATGCTCATAGGCGACGATTTGACTTGGGAGCGACTCTCT 720
QY 1042 CGATCTATTCATGAAGGAATTGCAACGCCATTATGAATAGGAGCTGTCTAGTCTTTGAC 1101
DB 721 CGATCTATTCATGAAGGAATTGCAACGCCATTATGAATAGGAGCTGTCTAGTCTTTGAC 780
QY 1102 ACCGACTTGCACAGCTTTTTCGCAAAATGGCAATTTAGGCATCAATGTAATTTTCC 1161
DB 781 ACCGACTTGCACAGCTTTTTCGCAAAATGGCAATTTAGGCATCAATGTAATTTTCC 840
QY 1162 ATGCTGTGAATGGCAATCAAAACATTTTCTGGCCAGTGTTTTAAACTTCAGTTTCACAGA 1221
DB 841 ATGCTGTGAATGGCAATCAAAACATTTTCTGGCCAGTGTTTTAAACTTCAGTTTCACAGA 900
QY 1222 AAATAGGCGACCATCTGCTGCCCACTTAAACTCTTTTCGGTAGGTGGAAGC 1274
DB 901 AAATAGGCGACCATCTGCTGCCCACTTAAACTCTTTTCGGTAGGTGGAAGC 953
```

Search completed: April 24, 2005, 20:02:09  
Job time : 767 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 19:16:33 ; Search time 5657 Seconds  
(without alignments)  
10912.488 Million cell updates/sec

Title: US-10-679-246-1  
Perfect score: 1274  
Sequence: 1 ttctcttggtgtttatggt.....ctcttcgtagtggaagc 1274

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sta:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1274	100.0	1274	6	AR411820 Sequence
2	1274	100.0	1274	6	AX058082 Sequence
3	1274	100.0	2924	6	CQ491092 Sequence
4	1274	100.0	2924	6	CQ493590 Sequence
5	1274	100.0	2924	6	CQ496955 Sequence
6	1274	100.0	31705	9	HS4400826
7	1274	100.0	173304	9	AC023818
8	1272.4	99.9	2829	6	AX833145
9	1272.4	99.9	2829	6	AK094663
10	1270.8	99.7	2972	9	HS4807215
11	1249	98.0	4090	6	AX780380
12	1104	86.7	2128	9	AK056051
13	1042.8	81.9	2440	6	AR380043
14	958	75.2	1540	9	BC035562
15	958	75.2	1874	6	CQ719559
16	958	75.2	2048	9	HSU76247
17	957	75.1	2454	9	BC042550
18	951.4	74.7	1886	9	HSU63295
19	947	74.3	2034	9	BC018193

20	928.4	72.9	1884	6	A63558	A63558 Sequence 11
21	928.4	72.9	1884	6	AR271267	AR271267 Sequence
22	836	65.6	175636	10	AC142211	AC142211 Mus muscu
23	825.4	64.8	220157	2	AC105462	AC105462 Rattus no
24	823.8	64.7	232382	2	AC098991	AC098991 Rattus no
25	807.8	63.4	1645	10	BC046317	BC046317 Mus muscu
26	807.8	63.4	1968	6	AX305803	AX305803 Sequence
27	807.8	63.4	1968	10	MMSIAH1B	Z19579 M.musculus
28	791.8	62.2	1713	10	MMSIAH1B	Z19580 M.musculus
29	791.8	62.2	112893	10	AL732294	AL732294 Mouse DNA
30	791.8	62.2	234105	10	AC091606	AC091606 Mus Muscu
31	791.6	62.1	1465	10	AF389476	AF389476 Rattus no
32	790.2	62.0	1720	10	BC052887	BC052887 Mus muscu
33	780.4	61.3	1457	10	AB067814	AB067814 Rattus no
34	755.8	59.3	194622	10	AC133654	AC133654 Mus muscu
35	674.6	53.0	168091	10	AC139553	AC139553 Mus muscu
36	674.6	53.0	182733	10	AC123941	AC123941 Mus muscu
37	651.4	51.1	1246	5	BC072747	BC072747 Xenopus l
38	649.8	51.0	89948	9	AL359734	AL359734 Human DNA
39	643.2	50.5	2123	10	MMSIAHPSA	Z19582 M.musculus
40	586.8	46.1	711	5	CR386772	CR386772 Gallus ga
41	540.6	42.4	120810	5	BX649644	BX649644 Zebrafish
42	539	42.3	241301	5	BX470163	BX470163 Zebrafish
43	537.8	42.2	2419	5	BC045870	BC045870 Danio rer
44	487.8	38.3	600	5	AF411976	AF411976 Gallus ga
45	429	33.7	975	6	CQ717814	CQ717814 Sequence

ALIGNMENTS

RESULT 1	AR411820	Sequence 1 from patent US 6638734.	1274 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	AR411820					
DEFINITION	AR411820					
ACCESSION	AR411820					
VERSION	AR411820.1	GI:40164258				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1274)					
AUTHORS	Reed, J.C. and Matsuzawa, S.-i.					
TITLE	Nucleic acid encoding proteins involved in protein degradation,					
	products and methods related thereto					
JOURNAL	Patent: US 6638734-A.1 28-OCT-2003;					
FEATURES	Location/Qualifiers					
source	1..1274					
	/organism="unknown"					
	/mol_type="genomic DNA"					

ORIGIN	Query Match	100.0%	Score 1274;	DB 6;	Length 1274;
	Best Local Similarity	100.0%	Pred. No. 6.2e-271;		
	Matches 1274;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTTCTTTAGTCTTTTATGGTCCATTTTCTATTTTAGCATTATTATTTCTATGAGTCTAT	60		
Db	1	TTTCTTTAGTCTTTTATGGTCCATTTTCTATTTTAGCATTATTATTTCTATGAGTCTAT	60		
QY	61	CCAAAGACGATTAAGGGAGTTCACATGTTTCCGGAACATTTTGAAGAAGAGCTTATC	120		
Db	61	CCAAAGACGATTAAGGGAGTTCACATGTTTCCGGAACATTTTGAAGAAGAGCTTATC	120		
QY	121	CAGTGTACAGATCCTTAATAAGTGACATTCAGTGTAAATTTTATTTTAAATCTCTTT	180		
Db	121	CAGTGTACAGATCCTTAATAAGTGACATTCAGTGTAAATTTTATTTTAAATCTCTTT	180		
QY	181	TTAATCCTATTTTCT	240		
Db	181	TTAATCCTATTTTCT	240		
QY	241	TATGGCATGTAACATTTATTTATAAGTAAGTCATGGTTATTAATTTTCTCTCTGCT	300		





```
QY 541 ACTTGGCGGGCCCTTTGGGATCCATTGCAACTTGGCTATGAGAAAGTGGCTAATTCA 600
DB 541 ACTTGGCGGGCCCTTTGGGATCCATTGCAACTTGGCTATGAGAAAGTGGCTAATTCA 600
QY 601 GTACTTTTCCCTGTAATAATATGCTCTCTGGATGTGAAATAACTCTGCCACACAGAA 660
DB 601 GTACTTTTCCCTGTAATAATATGCTCTCTGGATGTGAAATAACTCTGCCACACAGAA 660
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGGCCCTTATTCCTGTCGGTGGCTGTCT 720
DB 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGGCCCTTATTCCTGTCGGTGGCTGTCT 720
QY 721 TCTGTAAATGGAAGGCTCTCTGGATGTGTAATGCCCATCTGTATGATCAGCATAG 780
DB 721 TCTGTAAATGGAAGGCTCTCTGGATGTGTAATGCCCATCTGTATGATCAGCATAG 780
QY 781 TCCATTACAACTTACAGGAGAGGATATAGTTTCTTGTCTACACACATTAATCTTCT 840
DB 781 TCCATTACAACTTACAGGAGAGGATATAGTTTCTTGTCTACACACATTAATCTTCT 840
QY 841 GGTGCTGTTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTCATTTAGTCTTTA 900
DB 841 GGTGCTGTTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTCATTTAGTCTTTA 900
QY 901 GAGAAACAGGAAATAACGATGCTCAACGAGCTTCTTGGCAATCGTACAGCTGATG 960
DB 901 GAGAAACAGGAAATAACGATGCTCAACGAGCTTCTTGGCAATCGTACAGCTGATG 960
QY 961 ACACGCAAGCAAGCTCAAAATTTTGTCTACCGACTTGGCTAAATGCTATAGGCA 1020
DB 961 ACACGCAAGCAAGCTCAAAATTTTGTCTACCGACTTGGCTAAATGCTATAGGCA 1020
QY 1021 TTGACTTGGGAAGCGACTCTCTGATCTATTTCATGAAGGAATTGCAACAGCAATTATGAAT 1080
DB 1021 TTGACTTGGGAAGCGACTCTCTGATCTATTTCATGAAGGAATTGCAACAGCAATTATGAAT 1080
QY 1081 AGGACTGTCTAGTCTTTGACACAGCATGTCAGCTTTTGGCAAAATGCAAAATGCAATTTA 1140
DB 1081 AGGACTGTCTAGTCTTTGACACAGCATGTCAGCTTTTGGCAAAATGCAAAATGCAATTTA 1140
QY 1141 GGCATCAATGTAATTTTCCATGTTGTAATGGAATGGAATCAACATTTTGGCCAGTCT 1200
DB 1141 GGCATCAATGTAATTTTCCATGTTGTAATGGAATGGAATCAACATTTTGGCCAGTCT 1200
QY 1201 TTAATACTTCAAGTTCACAGAAATAAGCACCCTCTGCTGCCAACCTTAAACTCTTT 1260
DB 1201 TTAATACTTCAAGTTCACAGAAATAAGCACCCTCTGCTGCCAACCTTAAACTCTTT 1260
QY 1261 CGGTAGTGGAGC 1274
DB 1261 CGGTAGTGGAGC 1274
```

```
RESULT 3
LOCUS CO491092
DEFINITION Sequence 22959 from Patent WO0160860.
ACCESSION CO491092
VERSION CO491092.1 GI:41456711
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 22959 23-AUG-2001;
FEATURES Millennium Predictive Medicine, Inc. (US)
source Location/Qualifiers
1..2924
```

PAT 30-JAN-2004

/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

```
ORIGIN
Query Match 100.0%; Score 1274; DB 6; Length 2924;
Best Local Similarity 100.0%; Pred. No. 5.9e-271; Indels 0; Gaps 0;
Matches 1274; Conservative 0; Mismatches 0;

QY 1 TTTCTTTAGTGTGTTATGTTCCATTTTCTATTTTATTTAGCATTTATTTCTATGTAGTCTAT 60
DB TTTCTTTAGTGTGTTATGTTCCATTTTCTATTTTATTTAGCATTTATTTCTATGTAGTCTAT 539
QY 61 CCAAGACGATTAAGGGAGTTCACATGTTTCCGGAAACATTTTGAAGAGAGAGCTTATC 120
DB CCAAGACGATTAAGGGAGTTCACATGTTTCCGGAAACATTTTGAAGAGAGAGCTTATC 599
QY 121 CAGTGTACAGATCCCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT 180
DB CAGTGTACAGATCCCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT 659
QY 181 TTAATCCTATTTTCTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAACCTTTAAAGGACT 240
DB TTAATCCTATTTTCTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAACCTTTAAAGGACT 719
QY 241 TATGGCATGTAAACATTTATTAAGTAAGTCATGGTTATTAATTTTCTTCTGCTT 300
DB TATGGCATGTAAACATTTATTAAGTAAGTCATGGTTATTAATTTTCTTCTGCTT 779
QY 301 CCTTATGTTATTTTTCAGAAATGAGCCGTGACATGCTGACATGAGCATTACTACCGGTACC 360
DB CCTTATGTTATTTTTCAGAAATGAGCCGTGACATGCTGACATGAGCATTACTACCGGTACC 839
QY 361 TCGAAGTGTCCACCATCCCGAGAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB TCGAAGTGTCCACCATCCCGAGAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
QY 421 GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTTGTGACTATGTGTACCGCCATCTTT 480
DB GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTTGTGACTATGTGTACCGCCATCTTT 959
QY 481 CAATGTCAAGTGGCCATCTCTGTTAGCAACTGTGCGCCCAAGCTCAGTGTCTGCTCCA 540
DB CAATGTCAAGTGGCCATCTCTGTTGTAGCAACTGTGCGCCCAAGCTCAGTGTCTGCTCCA 1019
QY 541 ACTTGGCGGGCCCTTTGGGATCCATTGCAACTTGGCTATGAGAAAGTGGCTAATTCA 600
DB ACTTGGCGGGCCCTTTGGGATCCATTGCAACTTGGCTATGAGAAAGTGGCTAATTCA 1079
QY 601 GTACTTTTCCCTGTAATAATATGCTCTTCTGGATGTGAAATAACTCTGCCACACAGAA 660
DB GTACTTTTCCCTGTAATAATATGCTCTTCTGGATGTGAAATAACTCTGCCACACAGAA 1139
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGGCCCTTATTCCTGTCGGTGGCTGTCT 720
DB AAAGCAGACCATGAAGAGCTCTGTGAGTTTGGCCCTTATTCCTGTCGGTGGCTGTCT 1199
QY 721 TCTGTAAATGCAAGGCTCTCTGATGCTGTAAATGCCCATCTGTATGATCAGCATAG 780
DB TCTGTAAATGCAAGGCTCTCTGATGCTGTAAATGCCCATCTGTATGATCAGCATAG 1259
QY 781 TCCATTACAACTTACAGGAGAGGATATAGTTTCTTGTCTACAGACATTAATCTTCTCT 840
DB TCCATTACAACTTACAGGAGAGGATATAGTTTCTTGTCTACAGACATTAATCTTCTCT 1319
QY 841 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTGGCTTTTCACTTCACTGTAGTCTTA 900
DB GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTGGCTTTTCACTTCACTGTAGTCTTA 1379
QY 901 GAGAAACAGGAAATAACGATGCTCACCAGAGTCTTCTCGCAATCGTACAGCTGATGAG 960
DB GAGAAACAGGAAATAACGATGCTCACCAGAGTCTTCTCGCAATCGTACAGCTGATGAG 1439
```

QY 961 ACACCAAGCAAGCTGAAATTTTCTTACCGAATTGAGCTAAATGGTCTATAGCGGACGA 1020  
DB 1440 ACACCAAGCAAGCTGAAATTTTCTTACCGAATTGAGCTAAATGGTCTATAGCGGACGA 1499  
QY 1021 TTGACTTTGGGAGCGACTCCTCGATCTATCATGAAGCAATTTGCAACAGCCATTATGAAT 1080  
DB 1500 TTGACTTTGGGAGCGACTCCTCGATCTATCATGAAGCAATTTGCAACAGCCATTATGAAT 1559  
QY 1081 AGCGACTGTCTAGTCTTTTGACACCAAGCAATTTGCAACAGCCATTATGAATTTTGA 1140  
DB 1560 AGCGACTGTCTAGTCTTTTGACACCAAGCAATTTGCAACAGCCATTATGAATTTTGA 1619  
QY 1141 GGCATCAATGTAATTTTCCATGTTGCAATGCAATGCAATTTTCTGGCCAGTGT 1200  
DB 1620 GGCATCAATGTAATTTTCCATGTTGCAATGCAATGCAATTTTCTGGCCAGTGT 1679  
QY 1201 TTAATACTTCAGTTTCACAGAAATAAGGCAACCCATCTGTCTGCCAACCCTTAAACTCTTT 1260  
DB 1680 TTAATACTTCAGTTTCACAGAAATAAGGCAACCCATCTGTCTGCCAACCCTTAAACTCTTT 1739  
QY 1261 CGGTAGGTGGAAGC 1274  
DB 1740 CGGTAGGTGGAAGC 1753  
RESULT 4  
LOCUS CO493590 2924 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 25457 from Patent WO0160860.  
ACCESSION CO493590  
VERSION CO493590.1 GI:41459209  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 Schlegel, R., Endege, W.O. and Monahan, J.E.  
Genes differentially expressed in human prostate cancer and their  
use  
JOURNAL Patent: WO 0160860-A 25457 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..2924  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1274; DB 6; Length 2924;  
Best Local Similarity 100.0%; Pred. NO. 5.9e-271; Indels 0; Gaps 0;  
Matches 1274; Conservative 0; Mismatches 0;  
QY 1 TTTCTTTAGTTGTTATGGTCCATTTTCTATTTTAGCAATTTATTTCTATAGTCTAT 60  
DB 480 TTTCTTTAGTTGTTATGGTCCATTTTCTATTTTAGCAATTTATTTCTATAGTCTAT 539  
QY 61 CCAAGACGATTAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120  
DB 540 CCAAGACGATTAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 599  
QY 121 CAGTGTACAGATCTTAATGAAGTGCACATTCAGTGTAAATTTTATATATCTTTT 180  
DB 600 CAGTGTACAGATCTTAATGAAGTGCACATTCAGTGTAAATTTTATATATCTTTT 659  
QY 181 TTAATCCTATTTTCTTCTCTTTTGTCTAGTAAATTTTGTATGAATCTTTAAAGGACT 240  
DB 660 TTAATCCTATTTTCTTCTCTTTTGTCTAGTAAATTTTGTATGAATCTTTAAAGGACT 719  
QY 241 TATGGCATGTAACATTTATTAAGTGAAGTCAATGTTATATTTTCTCTGCTCCT 300  
DB 720 TATGGCATGTAACATTTATTAAGTGAAGTCAATGTTATATTTTCTCTGCTCCT 779

QY 301 CCTTATGTTATTTTTCAGAAATGAGCGGTGCTACAGACTGCTACAGATTACTACGGTACC 360  
DB 780 CCTTATGTTATTTTTCAGAAATGAGCGGTGCTACAGACTGCTACAGATTACTACGGTACC 839  
QY 361 TCGAAGTGTCCACATCCCAAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 840 TCGAAGTGTCCACATCCCAAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899  
QY 421 GACTTGGCGAGTCTTTTGGAGTTCAGTCTGCTTTCAGTATGTTTACGCTGCTGCTGCTGCT 480  
DB 900 GACTTGGCGAGTCTTTTGGAGTTCAGTCTGCTTTCAGTATGTTTACGCTGCTGCTGCTGCT 959  
QY 481 CAATGTCCAGAGTGCCTATCTTTTGTAGCAACTGTCGCCAAAGCTCACATGTTGTCCA 540  
DB 960 CAATGTCCAGAGTGCCTATCTTTTGTAGCAACTGTCGCCAAAGCTCACATGTTGTCCA 1019  
QY 541 ACTTGGCGGGCCCTTTGGGATCCATTTGCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 1020 ACTTGGCGGGCCCTTTGGGATCCATTTGCGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079  
QY 601 GTACTTTTCCCTGTAATATGCTTCTTCTGATGTGAAATAAACTCTGCCACACACAGAA 660  
DB 1080 GTACTTTTCCCTGTAATATGCTTCTTCTGATGTGAAATAAACTCTGCCACACACAGAA 1139  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTAGGCTTATTTCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 1140 AAAGCAGACCATGAAGAGCTCTGTGAGTTAGGCTTATTTCTGCTGCTGCTGCTGCTGCTGCT 1199  
QY 721 TCCTGTAATGCGAAGGCTCTGCTGATGCTTAATGCCCATCTGATGCTGATGCTGATGCTGATGCT 780  
DB 1200 TCCTGTAATGCGAAGGCTCTGCTGATGCTTAATGCCCATCTGATGCTGATGCTGATGCTGATGCT 1259  
QY 781 TCCATTACAACTTACAGGAGAGGATATAGTTTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
DB 1260 TCCATTACAACTTACAGGAGAGGATATAGTTTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1319  
QY 841 GGTGCTGTGAGTGGGTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
DB 1320 GGTGCTGTGAGTGGGTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1379  
QY 901 GAGAACAGGAAATACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
DB 1380 GAGAACAGGAAATACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439  
QY 961 ACACCAAGCAAGCTGAAATTTTCTTACCGAATTGAGCTAAATGGTCTATAGCGGACGA 1020  
DB 1440 ACACCAAGCAAGCTGAAATTTTCTTACCGAATTGAGCTAAATGGTCTATAGCGGACGA 1499  
QY 1021 TTGACTTTGGGAGCGACTCCTCGATCTATTCATGAAGCAATTTGCAACAGCCATTATGAAT 1080  
DB 1500 TTGACTTTGGGAGCGACTCCTCGATCTATTCATGAAGCAATTTGCAACAGCCATTATGAAT 1559  
QY 1081 AGCGACTGTCTAGTCTTTTGACACCAAGCAATTTGCAACAGCCATTATGAATTTTGA 1140  
DB 1560 AGCGACTGTCTAGTCTTTTGACACCAAGCAATTTGCAACAGCCATTATGAATTTTGA 1619  
QY 1141 GGCATCAATGTAATTTTCCATGTTGCAATGCAATGCAATTTTCTGGCCAGTGT 1200  
DB 1620 GGCATCAATGTAATTTTCCATGTTGCAATGCAATGCAATTTTCTGGCCAGTGT 1679  
QY 1201 TTAATACTTCAGTTTCACAGAAATAAGGCAACCCATCTGTCTGCCAACCCTTAAACTCTTT 1260  
DB 1680 TTAATACTTCAGTTTCACAGAAATAAGGCAACCCATCTGTCTGCCAACCCTTAAACTCTTT 1739  
QY 1261 CGGTAGGTGGAAGC 1274  
DB 1740 CGGTAGGTGGAAGC 1753  
RESULT 5  
LOCUS CO496955 2924 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 28822 from Patent WO0160860.



ACCESSION CQ496955  
VERSION CQ496955.1 GI:41462591  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.  
TITLE Genes differentially expressed in human prostate cancer and their use  
JOURNAL Patent: WO 0160860-A 28822 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1..2924  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1274; DB 6; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 5.9e-271;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 TTTCTTTAGTTGTTATGTCCTCAATTTCTATTTAGCATTTATTTCTATGCTAT 60  
480 TTTCTTTAGTTGTTATGTCCTCAATTTCTATTTAGCATTTATTTCTATGCTAT 539  
61 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACAATTTGAAAGAGAGCTTATC 120  
540 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACAATTTGAAAGAGAGCTTATC 599  
121 CAGGTGACAGATCCTAATAAGTGACATTCAGTGTAATTTTATTTTAAATATCTTTT 180  
600 CAGGTGACAGATCCTAATAAGTGACATTCAGTGTAATTTTATTTTAAATATCTTTT 659  
181 TTAATCCTATTTTCTCTCTTTTCTGTCAGTAAATTTTGTATGAAACTTTAAAGGACT 240  
660 TTAATCCTATTTTCTCTCTTTTCTGTCAGTAAATTTTGTATGAAACTTTAAAGGACT 719  
241 TATGCGATGTAACATTTATTTATAAGTAAGTCATGTTATATTTTCTCTGCTCT 300  
720 TATGCGATGTAACATTTATTTATAAGTAAGTCATGTTATATTTTCTCTGCTCT 779  
301 CCTTATGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 360  
780 CCTTATGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 839  
361 TCGAAGTGTCCACATCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
840 TCGAAGTGTCCACATCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899  
421 GACTTCGCGAGTCTTTTTCAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
900 GACTTCGCGAGTCTTTTTCAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959  
481 CAAATGTCAGAGTGGCCATCTTTTTCAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
960 CAAATGTCAGAGTGGCCATCTTTTTCAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019  
541 ACTTGGCGGGGCCCTTTTGGATCCATTCGCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
1020 ACTTGGCGGGGCCCTTTTGGATCCATTCGCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079  
601 GTACTTTTCCCTGTAATATGCTCTTCTGGATGTGAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
1080 GTACTTTTCCCTGTAATATGCTCTTCTGGATGTGAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCT 1139  
661 AAAGCAGACATGAAGAGCTCTGCTGAGTTTATGAGCTTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
1140 AAAGCAGACATGAAGAGCTCTGCTGAGTTTATGAGCTTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1199  
721 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCAATCTGATGATCAGCATAAG 780

Db 1200 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGATGCATCAGCATAAG 1259  
QY 781 TCCATTACAACCTCAGAGGAGAGGATATAGTTTTTCTGCTACAGACATTAATCTTCCT 840  
Db 1260 TCCATTACAACCTCAGAGGAGAGGATATAGTTTTTCTGCTACAGACATTAATCTTCCT 1319  
QY 841 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTGTTGGCTTTTCACATCATGTTAGTCTTA 900  
Db 1320 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTGTTGGCTTTTCACATCATGTTAGTCTTA 1379  
QY 901 GAGAACAGAAAATACGATGTCACACGAGCTTCTTCGCAATCGTACAGCTGATAGGA 960  
Db 1380 GAGAACAGAAAATACGATGTCACACGAGCTTCTTCGCAATCGTACAGCTGATAGGA 1439  
QY 961 ACACGCAAGCAAGCTGAAAATTTTGCTTACCGACTTTGAGCTTAATGCTCATGGCAGCA 1020  
Db 1440 ACACGCAAGCAAGCTGAAAATTTTGCTTACCGACTTTGAGCTTAATGCTCATGGCAGCA 1499  
QY 1021 TTGACTTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATTCGCAACAGCCATATGAAT 1080  
Db 1500 TTGACTTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATTCGCAACAGCCATATGAAT 1559  
QY 1081 AGCGACTGCTAGTCTTTGACACACGAGCTTGCACAGCTTTTTCGCAAAAATGCAATTTA 1140  
Db 1560 AGCGACTGCTAGTCTTTGACACACGAGCTTGCACAGCTTTTTCGCAAAAATGCAATTTA 1619  
QY 1141 GGCAATCAATGTAATTTTCCATGTTGCAATGGAATGGAATGGAATGGAATGGAATGGAAT 1200  
Db 1620 GGCAATCAATGTAATTTTCCATGTTGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1679  
QY 1201 TTAATACTTCAGTTTTCAGAGAAAATAGGACCCATCTCTGCTGCAACCTTAAACTCTTT 1260  
Db 1680 TTAATACTTCAGTTTTCAGAGAAAATAGGACCCATCTCTGCTGCAACCTTAAACTCTTT 1739  
QY 1261 CGGTAGTGGAGC 1274  
Db 1740 CGGTAGTGGAGC 1753

RESULT 6  
HS400626  
LOCUS Homo sapiens SIAH1 gene, exons 1-2. 31705 bp DNA linear PRI 02-APR-2001  
DEFINITION Homo sapiens SIAH1 gene, exons 1-2.  
ACCESSION AJ400626  
VERSION AJ400626.1 GI:13539602  
KEYWORDS siah1 gene.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Medhioub, M., Muchardt, C., Tubacher, E., Giudicelli, C.,  
Hors-Cayla, M.C. and Thomas, G.  
TITLE Down regulation of the TATA-less and GC-rich SIAH1 promoter by TP53  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 31705)  
AUTHORS Medhioub, M.  
TITLE Direct Submission  
JOURNAL Submitted (11-APR-2000) Medhioub M., Fondation Jean DAUSSET. CEPH,  
27, rue Juliette Dodu, Paris 75010, FRANCE  
FEATURES  
source Location/Qualifiers  
1..31705  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="16ql2-ql3"  
5126..30193  
/gene="SIAH1"  
5126..5550  
/gene="SIAH1"  
/number=1  
/evidence=experimental



JOURNAL Submitted (03-MAY-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 4 (bases 1 to 173304)  
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.  
TITLE Direct Submission  
JOURNAL Submitted (03-JAN-2004) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Jan 3, 2004 this sequence version replaced gi:20429291.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.  
FEATURES  
source  
1. .173304  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="CTD-2600H12"  
ORIGIN  
Query Match 100.0%; Score 1274; DB 9; Length 173304;  
Best Local Similarity 100.0%; Pred. No. 4,7e-271;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTTAGTTGTTATGTCCTATTTCTATTTAGCATTTATTTCTATGAGTCTAT 60  
DB 66802 TTTCTTTAGTTGTTATGTCCTATTTCTATTTAGCATTTATTTCTATGAGTCTAT 66861  
QY 61 CCAAGACGATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 120  
DB 66862 CCAAGACGATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 66921  
QY 121 CAGGTACAGATCCTAATAAGTGACATTCAGTGAATTTTATTTTAAATATCTTT 180  
DB 66922 CAGGTACAGATCCTAATAAGTGACATTCAGTGAATTTTATTTTAAATATCTTT 66981  
QY 181 TTAATCCTATTTTCT 240  
DB 66982 TTAATCCTATTTTCT 67041  
QY 241 TATGGCATGTAACATTAATTAAGTAAGTCAATGTTATATTAATTTTCTCTCTCTCT 300  
DB 67042 TATGGCATGTAACATTAATTAAGTAAGTCAATGTTATATTAATTTTCTCTCTCTCT 67101  
QY 301 CTTATGTTATTTTTCAGAAATGAGCGGTGACATGCTACAGCATTAATACCGGTACC 360  
DB 67102 CTTATGTTATTTTTCAGAAATGAGCGGTGACATGCTACAGCATTAATACCGGTACC 67161  
QY 361 TCGAAGTGTCCACATCCAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 67162 TCGAAGTGTCCACATCCAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 67221  
QY 421 GACTTCGGCAGTCTTTTGTAGTCTCCAGTCTGCTTTGACTATGTTGTTACCGCCATCTT 480  
DB 67222 GACTTCGGCAGTCTTTTGTAGTCTCCAGTCTGCTTTGACTATGTTGTTACCGCCATCTT 67281  
QY 481 CAATGTACAGTGGCCATCTTTGTTGTAGCAATGTCGCCCAAGTCAATGTTGTCCA 540  
DB 67282 CAATGTACAGTGGCCATCTTTGTTGTAGCAATGTCGCCCAAGTCAATGTTGTCCA 67341  
QY 541 ACTTGGCGGGCCCTTTGGATCCATTCGCACTTGGCTATGAGAGAGTGGCTAATTC 600  
DB 67342 ACTTGGCGGGCCCTTTGGATCCATTCGCACTTGGCTATGAGAGAGTGGCTAATTC 67401  
QY 601 GTACTTTTCCCTGTAAATATGCTTCTGGAATGTAATTAATCTCTGCCACACAGAA 660  
DB 67402 GTACTTTTCCCTGTAAATATGCTTCTGGAATGTAATTAATCTCTGCCACACAGAA 67461

QY 661 AAAGCAGACCATCAAGAGCTCTGAGTTAGGCTTATTCCTGTCGTCCTGCTGCT 720  
DB 67462 AAAGCAGACCATCAAGAGCTCTGAGTTAGGCTTATTCCTGTCGTCCTGCTGCT 67521  
QY 721 TCTGTAAATGGCAAGGCTCTCTGGAATGCTGAATGCCCATCTGATGCAATCAGCATAAG 780  
DB 67522 TCTGTAAATGGCAAGGCTCTCTGGAATGCTGAATGCCCATCTGATGCAATCAGCATAAG 67581  
QY 781 TCCATTACAAACCTTACAGGAGAGGATATAGTTTTTCTTGTCTACAGACATTAATCTTCT 840  
DB 67582 TCCATTACAAACCTTACAGGAGAGGATATAGTTTTTCTTGTCTACAGACATTAATCTTCT 67641  
QY 841 GGTGCTGTTGACTGGGTGATGATGCTGCTGTTTGGCTTTCACCTTCATGTTAGTCTTA 900  
DB 67642 GGTGCTGTTGACTGGGTGATGATGCTGCTGTTTGGCTTTCACCTTCATGTTAGTCTTA 67701  
QY 901 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTTCGCAATCGTACAGCTGATAGGA 960  
DB 67702 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTTCGCAATCGTACAGCTGATAGGA 67761  
QY 961 ACACGCAAGCAAGCTGAAATTTTGTCTTACCGAATTTGAGCTAAATGCTCAAGCCATTAAT 1020  
DB 67762 ACACGCAAGCAAGCTGAAATTTTGTCTTACCGAATTTGAGCTAAATGCTCAAGCCATTAAT 67821  
QY 1021 TTGACTTGGGAAGCGACTCTCGATCTATTCATGAAGGAATTCGAAGCCATTAATGAAT 1080  
DB 67822 TTGACTTGGGAAGCGACTCTCGATCTATTCATGAAGGAATTCGAAGCCATTAATGAAT 67881  
QY 1081 AGCGACTGTCTAGTCTTTGACACACGATTTGACAGCTTTTTCGCAAAAATGCAATTTA 1140  
DB 67882 AGCGACTGTCTAGTCTTTGACACACGATTTGACAGCTTTTTCGCAAAAATGCAATTTA 67941  
QY 1141 GGCAATCAATGTAATTTTCCATGTTGAAATGGCAATCAAAATTTTCTGCGCCAGTGT 1200  
DB 67942 GGCAATCAATGTAATTTTCCATGTTGAAATGGCAATCAAAATTTTCTGCGCCAGTGT 68001  
QY 1201 TTAATACTTCAGTTTCACAGAAAATGAAGGACCCATCTGTCTGCCAACCTTAAACTCTTT 1260  
DB 68002 TTAATACTTCAGTTTCACAGAAAATGAAGGACCCATCTGTCTGCCAACCTTAAACTCTTT 68061  
QY 1261 CGTAGTGTGGAAGC 1274  
DB 68062 CGTAGTGTGGAAGC 68075  
RESULT 8  
AX833145  
LOCUS AX833145 2829 bp DNA linear PAT 15-DEC-2003  
DEFINITION Sequence 269 from Patent EPI347046.  
ACCESSION AX833145  
VERSION AX833145.1 GI:39919280  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, Y.  
TITLE Full-length cDNA sequences  
JOURNAL Patent: EP 1347046-A 269 24-SEP-2003;  
Research Association for Biotechnology (JP)  
FEATURES  
source  
1. .2829  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 99.9%; Score 1272.4; DB 6; Length 2829;  
Best Local Similarity 99.9%; Pred. No. 1.3e-270;



Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.

NEDO human cDNA sequencing project

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

3 (bases 1 to 2829)  
Isogai,T. and Yamamoto,J.  
Direct Submission  
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp; Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.

## FEATURES

## source

1..2829  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BRAMY2021139"  
/tissue\_type="anygdala"  
/clone\_lib="BRAMY2"  
/note="cloning vector: pME18SFLJ"

## ORIGIN

Query Match 99.9%; Score 1272.4; DB 9; Length 2829;  
Best Local Similarity 99.9%; Pred. No. 1.3e-270;  
Matches 1273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 TTTCTTTAGTTGTTATGTTCCATTTTCTATTATTAGCATTTATTATTCTATGATGCTAT 60  
1134 TTTCTTTAGTTGTTATGTTCCATTTTCTATTATTAGCATTTATTATTCTATGATGCTAT 1193  
61 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAAATTTTGAAGAGAGAGCTTATC 120  
1194 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAAATTTTGAAGAGAGAGCTTATC 1253  
121 CAGGTACAGATCCTAATAAGTGACATTCAGTGAATTTTATTATTATTATTATATCTTTT 180  
1254 CAGGTACAGATCCTAATAAGTGACATTCAGTGAATTTTATTATTATTATTATTATTATCTTT 1313  
181 TTAATCCTATTTTCTTCTCTTTTCTCAGTAAATTTTGTATGAACTTTTAAAGGACT 240  
1314 TTAATCCTATTTTCTTCTCTTTTCTCAGTAAATTTTGTATGAACTTTTAAAGGACT 1373  
241 TATGGCATGTAACATTTATATAAGTAAGTCATCGTTTATATTTTCTCTGCT 300  
1374 TATGGCATGTAACATTTATATAAGTAAGTCATCGTTTATATTTTCTCTGCT 1433  
301 CCTATGTTATTTATTTACAGAAATGAGCGCTCAGCTGCTACAGCATTTACCGGTACC 360  
1434 CCTATGTTATTTATTTACAGAAATGAGCGCTCAGCTGCTACAGCATTTACCGGTACC 1493  
361 TCGAAGTGTCCACATCCACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
1494 TCGAAGTGTCCACATCCACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553  
421 GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCCAATCTT 480  
1554 GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCCAATCTT 1613  
481 CAATGTACAGTGGCCATCTTTTGTAGCAATCTGCGCCCAAGCTTCATGTGTGCTCA 540  
1614 CAATGTACAGTGGCCATCTTTTGTAGCAATCTGCGCCCAAGCTTCATGTGTGCTCA 1673  
541 ACTTGGCGGGGCTTTTGGATGCTCATTTCCCACTTGGCTATGTGAGAAAGTGGCTAATCCA 600  
1674 ACTTGGCGGGGCTTTTGGATGCTCATTTCCCACTTGGCTATGTGAGAAAGTGGCTAATCCA 1733  
601 GTACTTTTCCCTGTAAATATATGCTCTTCTGGATGTGAAATAACTCTGCCACACAGAA 660

1734 GTACTTTTCCCTGTAAATATGCTCTTCTGGATGTGAAATAACTCTGCCACACAGAA 1793  
661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTATAGCCCTTATTTCTGTCCTGTCCTGGTCT 720  
1794 AAAGCAGACCATGAAGAGCTCTGTGAGTTTATAGCCCTTATTTCTGTCCTGTCCTGGTCT 1853  
721 TCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGATCAGCATAAG 780  
1854 TCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGATCAGCATAAG 1913  
781 TCCTATTACACCTTACAGGAGAGGATATAGTTTCTTGGCTACACACATTAATCTTCTCT 840  
1914 TCCTATTACACCTTACAGGAGAGGATATAGTTTCTTGGCTACACACATTAATCTTCTCT 1973  
841 CGTGCTGTTGACTGGGTGATGATGCACTGCTGTTTGGGCTTTCACATTCATGTTAGTCTTA 900  
1974 GGTGCTGTTGACTGGGTGATGATGCACTGCTGTTTGGGCTTTCACATTCATGTTAGTCTTA 2033  
901 GAGAAACAGGAAATAACGATGCTCACAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 960  
2034 GAGAAACAGGAAATAACGATGCTCACAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 2093  
961 ACACGACAGCAAGCTGAAATTTTGTCTTACCGACTTGAAGCTTAAATGCTCATAGGCCACGA 1020  
2094 ACACGACAGCAAGCTGAAATTTTGTCTTACCGACTTGAAGCTTAAATGCTCATAGGCCACGA 2153  
1021 TTGACTTTGGGAAGCGACTCTCGATCTATTTCATGAGGAATTCACACAGCCATTTATGAAT 1080  
2154 TTGACTTTGGGAAGCGACTCTCGATCTATTTCATGAGGAATTCACACAGCCATTTATGAAT 2213  
1081 AGCGACTGTCTAGTCTTTTGACACACAGCAATTCACAGCTTTTTCGAGAAATGCGCAATTTA 1140  
2214 AGCGACTGTCTAGTCTTTTGACACACAGCAATTCACAGCTTTTTCGAGAAATGCGCAATTTA 2273  
1141 GGATCAATGTAATCTATTTCATGTTGAAATGCGCAATTCATTTCTGSCCAGTGT 1200  
2274 GGATCAATGTAATCTATTTCATGTTGAAATGCGCAATTCATTTCTGSCCAGTGT 2333  
1201 TTAATACTTCAGTTTCACAGAAATAAGGACCCATCTCTGTCGCAACCTTAAACCTCTTT 1260  
2334 TTAATACTTCAGTTTCACAGAAATAAGGACCCATCTCTGTCGCAACCTTAAACCTCTTT 2393  
1261 CGGTAGGTGGAAGC 1274  
2394 CGGTAGGTGGAAGC 2407

## RESULT 10

## HSM807215

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## CONSTRM

## TITLE

## JOURNAL

## COMMENT

HSM807215 2972 bp mRNA linear PRI 30-AUG-2003  
Homo sapiens mRNA; cDNA DKFZp686L1897 (from clone DKFZp686L1897);  
complete cds.  
BX647064  
BX647064.1 GI:34366097  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2972)  
Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,  
Fobo,G., Han,M. and Wiemann,S.  
The German Human cDNA Consortium  
Direct Submission  
Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Sequenced by BMFZ (Biomedical Research Center at the  
Heinrich-Heine-University, Duesseeldorf/Germany) within the cDNA  
sequencing consortium of the German Genome Project. This clone  
(DKFZp686L1897) is available at the RZPD in Berlin. Please contact

the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

# FEATURES

Location/Qualifiers  
 1..2972  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="16q12.1"  
 /clone="DKFZp686L1897"  
 /tissue\_type="human retina"  
 /clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host DH10B; sites SfiIA + SfiIB"  
 /dev\_stage="adult"  
 1..2972  
 /gene="DKFZp686L1897"  
 1062..1910  
 /gene="DKFZp686L1897"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CAE46191.1"  
 /db\_xref="GI:34366098"  
 /translation="MSRQTATALPTGTSCPPSPQRPALPTGTASNNDLASLPECPVC FVVLPIIQCOSGHLVNCNRPKLTCCPTCRPLGSRINLAMEKVNDSVLPFPKYAS SCGLTLPTKADHEELCEFPYPCPGKSGWGLDAMPHLMHOKHSITTLQO EDIVFLATINLSGAVDWMQSCPGFHEMLVLEKQEKYDGHQOFAIVOLIGTRKOA ENFAYKLEINLGHRRRLTWEATPSRHGGIATAIMSDCLVFTSIAQLFAENGNLGIN VTISMG"  
 2931..2936  
 /polyA\_signal  
 2948  
 /polyA\_site  
 /gene="DKFZp686L1897"

# ORIGIN

Query Match 99.7%; Score 1270.8; DB 9; Length 2972;  
 Best Local Similarity 99.8%; Pred. No. 3e-270;  
 Matches 1272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 TTTCTTTAGTGTATGGTCCATTTTCTATTTTACGATTTATTTCTATGTAGTCTAT 60  
 741 TTTCTTTAGTGTATGGTCCATTTTCTATTTTACGATTTATTTCTATGTAGTCTAT 800  
 61 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 120  
 801 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 860  
 121 CAGTGTAAGATCTAATAAGTGCAATTCAGTGTAAATTTTATTTTATATCTTTT 180  
 861 CAGTGTAAGATCTAATAAGTGCAATTCAGTGTAAATTTTATTTTATATCTTTT 920  
 181 TTAATCTATTTTCTCTCTTTTCTGCTCAATAATTTTCTATGAAGCTTTAAAGGACT 240  
 921 TTAATCTATTTTCTCTCTTTTCTGCTCAATAATTTTCTATGAAGCTTTAAAGGACT 980  
 241 TATGGCATGTAACATTTATTAAGTAAAGTCAATGTTATATTTTCTCTGCTCT 300  
 981 TATGGCATGTAACATTTATTAAGTAAAGTCAATGTTATATTTTCTCTGCTCT 1040  
 301 CCTTATGATTTATTTTCAAGAAATGAGCCGTGAGCTGCTACAGATTTACCTACGGTACC 360  
 1041 CCTTATGATTTATTTTCAAGAAATGAGCCGTGAGCTGCTACAGATTTACCTACGGTACC 1100  
 361 TCGAGTGTCCACATCCAGAGGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 1101 TCGAGTGTCCACATCCAGAGGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1160  
 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTGTGACTATGTTTACCGCCCATCTTT 480  
 1161 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTGTGACTATGTTTACCGCCCATCTTT 1220  
 481 CAATGTGAGTGGCCATCTTGTGTTGTAGCACTGTGCGCCCAAGCTCACATGTTGCCA 540

Db 1221 CAATGTGAGTGGCCATCTTTGTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCA 1280  
 Qy 541 ACTTGGCGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATTC 600  
 Db 1281 ACTTGGCGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATTC 1340  
 Qy 601 GTACTTTTCCCTCTAATAATATGCGCTCTTCTGGATGTGAATAACTCTGCCACACAGAA 660  
 Db 1341 GTACTTTTCCCTCTAATAATATGCGCTCTTCTGGATGTGAATAACTCTGCCACACAGAA 1400  
 Qy 661 AAAGCAGACATGAGAGCTCTGTGAGTTTAGGCTTATTCCTGTCCTGCGCTGGTCT 720  
 Db 1401 AAAGCAGACATGAGAGCTCTGTGAGTTTAGGCTTATTCCTGTCCTGCGCTGGTCT 1460  
 Qy 721 TCCTGTAATGCGCAAGGCTCTCTGATGCTCTAATGCGCCCATCTGATGATCAGCATTAAG 780  
 Db 1461 TCCTGTAATGCGCAAGGCTCTCTGATGCTCTAATGCGCCCATCTGATGATCAGCATTAAG 1520  
 Qy 781 TCCATTACAAACCTTACAGGAGAGATATAGTTTCTTGTGTACAGCATTAATCTTCT 840  
 Db 1521 TCCATTACAAACCTTACAGGAGAGATATAGTTTCTTGTGTACAGCATTAATCTTCT 1580  
 Qy 841 GGTGCTGTGACTGGGTGATGATGAGTCCCTGTTTGGCTTTCCTTCACTTGTAGTCTTA 900  
 Db 1581 GGTGCTGTGACTGGGTGATGATGAGTCCCTGTTTGGCTTTCCTTCACTTGTAGTCTTA 1640  
 Qy 901 GAGAAACAGGAAAAATACGATGGTCCACAGCAGTTCCTCGCAATCGTACAGCTGATAGA 960  
 Db 1641 GAGAAACAGGAAAAATACGATGGTCCACAGCAGTTCCTCGCAATCGTACAGCTGATAGA 1700  
 Qy 961 ACACGCAAGAGCTGAAAAATTTTCTTACCGACTTGGCTTAAATGGTTCATAGGCGAGCA 1020  
 Db 1701 ACACGCAAGAGCTGAAAAATTTTCTTACCGACTTGGCTTAAATGGTTCATAGGCGAGCA 1760  
 Qy 1021 TTGACTTGGAGGAGCTCCTCGATCTATTATGAAGGAATTCACAGCCATTTATGAAT 1080  
 Db 1761 TTGACTTGGAGGAGCTCCTCGATCTATTATGAAGGAATTCACAGCCATTTATGAAT 1820  
 Qy 1081 AGCGACTGTCTAGTCTTTGACACAGCATTGACAGCTTTTGGCAGAAAAATGGCAATTTA 1140  
 Db 1821 AGCGACTGTCTAGTCTTTGACACAGCATTGACAGCTTTTGGCAGAAAAATGGCAATTTA 1880  
 Qy 1141 GGCATCAATGTAATTTTCCATGTTTGAATGCAATTCGAATTTTCTGCGCAGTCT 1200  
 Db 1881 GGCATCAATGTAATTTTCCATGTTTGAATGCAATTCGAATTTTCTGCGCAGTCT 1940  
 Qy 1201 TTAATACTTCACTTTCACAGAAAAATAGGCAACCCATCTCTGCGCAACCTTAAACTCTTT 1260  
 Db 1941 TTAATACTTCACTTTCACAGAAAAATAGGCAACCCATCTCTGCGCAACCTTAAACTCTTT 2000  
 Qy 1261 CGGTAGTGGAGC 1274  
 Db 2001 CGGTAGTGGAGC 2014

RESULT 11  
 AX780380/c 4090 bp DNA linear PAT 14-JUL-2003  
 LOCUS Sequence 2537 from Patent WO03039443.  
 DEFINITION AX780380  
 ACCESSION AX780380  
 VERSION AX780380.1 GI:32697374  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Haerlach, T.; Schoch, C.; Kern, W.; Kohlmann, A.; Schnittger, S.,  
 Dugas, M.; Ellis, R.; Brors, B. and Mergenthaler, S.  
 TITLE Novel genetic markers for leukemias  
 JOURNAL Patent: WO 03039443-A 2537 15-MAY-2003;  
 Deutsches Krebsforschungszentrum (DKFZ);  
 Ludwig-Maximilians-Universitaet Muenchen (LMU); Haerlach, Torsten,



FEATURES		PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)	
source	location/Qualifiers		
	1..4090		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	98.0% ; Score 1249 ; DB 6 ; Length 4090 ;		
Best Local Similarity	99.6% ; Pred. No. 1.9e-265 ;		
Matches 1271 ;	Conservative 0 ; Mismatches 3 ; Indels 2 ; Gaps 2 ;		
QY	1	TTTCTTTAGTTGTTATGGTCATTTCTATTTTAGCAATTTATTTCTATGATGCTAT	60
DB	3611	TTTCTTTAGTTGTTATGGTCATTTCTATTTTAGCAATTTATTTCTATGATGCTAT	3552
QY	61	CAAAGACGATTAAAGGAGTTCCATGTTTTCGGAACATTTTGAAGAGAGGCTTATC	120
DB	3551	CAAAGACGATTAAAGGAGTTCCATGTTTTCGGAACATTTTGAAGAGAGGCTTATC	3492
QY	121	CAGTGACAGATCCTTAATAAGTGACATTCAGTGTAATTTTATTTTAAATATCTTTT	180
DB	3491	CAGTGACAGATCCTTAATAAGTGACATTCAGTGTAATTTTATTTTAAATATCTTTT	3432
QY	181	TTAATCCTATTTTCTCTCTCTTTGCTCAGTAAATTTTGTATGAAGGACT	240
DB	3431	TTAATCCTATTTTCTCTCTCTTTGCTCAGTAAATTTTGTATGAAGGACT	3372
QY	241	TATGGCATGTAACATTATTTATAAGTAAGTCATGGTTATTAATTTTCTCTGCCT	300
DB	3371	TATGGCATGTAACATTATTTATAAGTAAGTCATGGTTATTAATTTTCTCTGCCT	3312
QY	301	CCTTATGTTATTTTTCAGAAATGAGCGTCAGACTGCTACAGCAATTCACCGGTACC	360
DB	3311	CCTTATGTTATTTTTCAGAAATGAGCGTCAGACTGCTACAGCAATTCACCGGTACC	3252
QY	361	TGCAAGTGTCCACATCCAGAGGGTGTGCGCTGACTGGGCAACTGCATCAACAAT	420
DB	3251	TGCAAGTGTCCACATCCAGAGGGTGTGCGCTGACTGGGCAACTGCATCAACAAT	3192
QY	421	GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTACACTATGTGTTACCGCCATCTT	480
DB	3191	GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTACACTATGTGTTACCGCCATCTT	3132
QY	481	CAATGTCCAGATGGCCATCTTTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCA	540
DB	3131	CAATGTCCAGATGGCCATCTTTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCA	3072
QY	541	ACTTCCGGGGCCCTTTTGGATCCATTCGCAACTTGGCTATGGAGAAAGTGCTAATCCA	600
DB	3071	ACTTCCGGGGCCCTTTTGGATCCATTCGCAACTTGGCTATGGAGAAAGTGCTAATCCA	3012
QY	601	GTACTTTTCCCTGTAAATATGCTCTTCTGGATGTGAATTAACCTGCGCACACAGAA	660
DB	3011	GTACTTTTCCCTGTAAATATGCTCTTCTGGATGTGAATTAACCTGCGCACACAGAA	2952
QY	661	AAA-GCAGACCACTGAAGA-GCTCTGTGAGTTTAGGCTTTATTCCTGTCGGTCCCTGGTG	718
DB	2951	AAANGCAGACCACTGAANANGCTCTGTGAGTTTAGGCTTTATTCCTGTCGGTCCCTGGTG	2892
QY	719	CTTCTGTTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGATCATGACATA	778
DB	2891	CTTCTGTTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGATCATGACATA	2832
QY	779	AGTCATTTACACCCCTACAGGAGGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTC	838
DB	2831	AGTCATTTACACCCCTACAGGAGGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTC	2772
QY	839	CTGGTGCTGTTGACTGGGTGATGATGCTGCTCTGTTTGGCTTTTCACTTCACTGTTAGTCT	898
DB	2771	CTGGTGCTGTTGACTGGGTGATGATGCTGCTCTGTTTGGCTTTTCACTTCACTGTTAGTCT	2712
QY	899	TAGAGAAAACAGGAAAAATACGATGGTCCACGAGCAAGTTCTTTCGCAATCGTACAGCTGATAG	958

DB	2711	TAGAGAAAACAGGAAAAATACGATGGTCCACGAGTTCTTCGCAATCGTACAGCTGATAG	2652
QY	959	GAACACGCAAGCAAGCTGAAAAATTTTCTTACCGACTTTCAGCTTAATGGTCTATAGCGCAC	1018
DB	2651	GAACACGCAAGCAAGCTGAAAAATTTTCTTACCGACTTTCAGCTTAATGGTCTATAGCGCAC	2592
QY	1019	GATTGACTTTGGGAAGCACTCTCGATCTATTTCATGAAGAAATTCGCAAGCATTATGA	1078
DB	2591	GATTGACTTTGGGAAGCACTCTCGATCTATTTCATGAAGAAATTCGCAAGCATTATGA	2532
QY	1079	ATAGGACGCTCTAGCTCTTTGACACGACGATTCGACAGCTTTTTCGAGAAATGGCAATT	1138
DB	2531	ATAGGACGCTCTAGCTCTTTGACACGACGATTCGACAGCTTTTTCGAGAAATGGCAATT	2472
QY	1139	TAGGCATCAATGTAATCTATTTCATGTTGTAATGGCAATCAAAACATTTTTCGGCCAGT	1198
DB	2471	TAGGCATCAATGTAATCTATTTCATGTTGTAATGGCAATCAAAACATTTTTCGGCCAGT	2412
QY	1199	GTTTAAAACTTCAGTTTCACAGAAATAAGGACCCATCTGCTGCCAACCTAAAACTCT	1258
DB	2411	GTTTAAAACTTCAGTTTCACAGAAATAAGGACCCATCTGCTGCCAACCTAAAACTCT	2352
QY	1259	TTTCGCTAGTGGAGC 1274	
DB	2351	TTTCGCTAGTGGAGC 2336	
RESULT 12			
AK056051/c			
LOCUS			
DEFINITION Homo sapiens cDNA FLJ31489 fis, clone NT2NE2003308.			
ACCESSION AK056051			
VERSION AK056051.1 GI:16551141			
KEYWORDS oligo capping; fis (full insert sequence).			
SOURCE Homo sapiens			
ORGANISM Homo sapiens			
REFERENCE 1			
AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohyashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahori, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshioka, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, P., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirose, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Tachibana, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.			
Complete sequencing and characterization of 21,243 full-length human cDNAs			
TITLE Nat. Genet. 36 (1), 40-45 (2004)			
JOURNAL			
PUBMED 14702039			

[illegible]



QY 229 TTTTAAAGGACTTATGCGCATGTAAACATTTATTTATTAAGTAAGTCATGTTTAATTAATT 288  
Db |||||  
QY 462 TTTTGAAGGACTTATGCGCATGTAAACATTTATTTATAAGTAAGTCATGTTTAATTAATT 521  
Db |||||  
QY 289 TTTTCTCTGCTCCTTATGTAATTTATTTTTCAGAAATGAGCGGTGCTGAGAGCTGCTACAGCATTA 348  
Db |||||  
QY 522 TTTTCTCTGCTCCTTATGTAATTTATTTTTCAGAAATGAGCGGTGCTGAGAGCTGCTACAGCATTA 581  
QY 349 CTTACCGGTACCTCGAAGTGTCCACATCCAGAGGGTGCTGCCCTGAGCTGGCAACAAT 408  
Db |||||  
QY 582 CTTACCGGTACCTCGAAGTGTCCACATCCAGAGGGTGCTGCCCTGAGCTGGCAACAAT 641  
QY 409 GCATCCAAACATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTGTGACTATGTGTTA 468  
Db |||||  
QY 642 GCATCCAAACATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTGTGACTATGTGTTA 701  
QY 469 CCGCCCATCTCTCAATGTGACAGTGGCCATCTTTTGTAGCAACTGTGCCCAAGCTC 528  
Db |||||  
QY 702 CCGCCCATCTCTCAATGTGACAGTGGCCATCTTTTGTAGCAACTGTGCCCAAGCTC 761  
QY 529 ACATGTGTCCAACTTGCOCGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAAA 588  
Db |||||  
QY 762 ACATGTGTCCAACTTGCOCGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAAA 821  
QY 589 GTGGCTAATTCAGTACTTTTCCCTGTAAATATGCGTCTCTGGATGTGAATAAATCTCTG 648  
Db |||||  
QY 822 GTGGCTAATTCAGTACTTTTCCCTGTAAATATGCGTCTCTGGATGTGAATAAATCTCTG 881  
QY 649 CCACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTTGAAGCTTATTCCTGTCG 708  
Db |||||  
QY 882 CCACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTTGAAGCTTATTCCTGTCG 941  
QY 709 TGCCTCTGCTTCTGTAAATGGCAAGCTCTCTGGATGCTGTAAATGCCCATCTGATG 768  
Db |||||  
QY 942 TGCCTCTGCTTCTGTAAATGGCAAGCTCTCTGGATGCTGTAAATGCCCATCTGATG 1001  
QY 769 CATCAGCATAACTCAATTAACCCCTACAGGAGAGGATATAGTTTTCTTGCTACAGAC 828  
Db |||||  
QY 1002 CATCAGCATAACTCAATTAACCCCTACAGGAGAGGATATAGTTTTCTTGCTACAGAC 1061  
QY 829 ATTAACTTCTCGTGCTGTGATGCTGGTGATGATGAGTCTGCTGTTTGGCTTTCACCTC 888  
Db |||||  
QY 1062 ATTAACTTCTCGTGCTGTGATGCTGGTGATGATGAGTCTGCTGTTTGGCTTTCACCTC 1121  
QY 889 ATGTTAGTCTTAGAGAAAACAGGAAAATACGATGTCACAGCAGTCTTTCGCAATCGTA 948  
Db |||||  
QY 1122 ATGTTAGTCTTAGAGAAAACAGGAAAATACGATGTCACAGCAGTCTTTCGCAATCGTA 1181  
QY 949 CAGCTGATAGGAAACAGCAAGCAAGCTGAAAATTTTGTCTTACGACTTGAGCTAAATGCT 1008  
Db |||||  
QY 1182 CAGCTGATAGGAAACAGCAAGCAAGCTGAAAATTTTGTCTTACGACTTGAGCTAAATGCT 1241  
QY 1009 CATAGCGAGATTGACTTGGGAGGAGCTCTCGATCTATTATCAAGGAATTCGAACA 1068  
Db |||||  
QY 1242 CATAGCGAGATTGACTTGGGAGGAGCTCTCGATCTATTATCAAGGAATTCGAACA 1301  
QY 1069 GCATTTATCAATAGGCACTGTCTAGTCTTTTGACACAGCAGTTCACAGCTTTTTCAGAAA 1128  
Db |||||  
QY 1302 GCATTTATCAATAGGCACTGTCTAGTCTTTTGACACAGCAGTTCACAGCTTTTTCAGAAA 1361  
QY 1129 AATGGCAATTTAGGCACTCAATGTAATTTTCAATGTGTGAATGGCAATCAAAATTT 1188  
Db |||||  
QY 1362 AATGGCAATTTAGGCACTCAATGTAATTTTCAATGTGTGAATGGCAATCAAAATTT 1421  
QY 1189 TCTGGCCAGTGTAAACCTCAGTTTTCAGAAATAGGCAATCCCATCTGCTGCGCAAC 1248  
Db |||||  
QY 1422 TCTGGCCAGTGTAAACCTCAGTTTTCAGAAATAGGCAATCCCATCTGCTGCGCAAC 1481  
QY 1249 CTAATAACTCTTTCGGTGTGGAGC 1274  
Db |||||  
QY 1482 CTAATAACTCTTTCGGTGTGGAGC 1507

RESULT 14  
BC035562  
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBLISHED

REFERENCE

AUTHORS

JOURNAL

REMARK

COMMENT

FEATURES  
source

BC035562 1540 bp mRNA linear PRI 02-JAN-2004  
Homo sapiens seven in absentia homolog 1 (Drosophila), mRNA (CDNA  
clone MGC:40343 IMAGE:5224755), complete cds.

BC035562  
BC035562.1 GI:23274141  
MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1540)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.P., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Fahsey,J., Helton,B., Kettman,M., Madan,A., Rodrigues,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Boutard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,

Schmerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1540)

Strausberg,R.

Direct Submission

Submitted (31-JUL-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgc@hgrl.nih.gov](mailto:nisc_mgc@hgrl.nih.gov)

Akter,N., Ayie,K., Beckstrom-Sternberg,S.M., Benjamin,B.,

Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,

Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,

Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,

Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,

McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,

Tsurgeon,C., Vogt,J.D., Walker,M.A., Wetherby,K.D., Wiggins,L.,

Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAP Plate: 64 Row: 0 Column: 8

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 4506946.

Location/Qualifiers

1. 1540

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:40343 IMAGE:5224755"

/tissue\_type="Pancreas, Spleen, adult pooled"

/clone\_lib="NIH\_MGC\_120"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

gene

1. 1540

/gene="STAH1"

/db\_xref="LOCUSID:6477"

/db\_xref="MIM:602212"

118. 966

/gene="STAH1"

/codon\_start=1

/product="STAH1 protein"

/protein\_id="AAH35562.1"

/db\_xref="GI:23274142"

/db\_xref="LOCUSID:6477"

/db\_xref="MIM:602212"

/translation="MSRQATATLPTGSKPSPQRPALVTGTTASNLDLSEFCPVC  
FDVLPPILOQSGHLVSGNCRPKLTCCPTCRGLPSIRNLAMKXVANSVLPFCVKYAS  
SGCEILPHEKADHEELCEPYPSPCPGSKWQSGSLDAVPHLMHOKHSITVLOG  
EDIVFLATINDLPADVWVMQSCFPHMLVLEKQEKYDGHQOQFALVOLIGTEKQA  
ENFAYRLNGLHRRRLTWEATPRSIEGFIATAMNSDCLVFDTSIAQLFAENGLNGLIN  
VTISMC"

361. 963

/gene="STAH1"

/note="Sina; Region: Seven in absentia protein family. The seven in absentia (sina) gene was first identified in Drosophila. The Drosophila Sina protein is essential for the determination of the R7 pathway in photoreceptor cell development; the loss of functional Sina results in the transformation of the R7 precursor cell to a non-neuronal cell type. The Sina protein contains an N-terminal RING finger domain Pfam00097. Through this domain, Sina binds E2 ubiquitin-conjugating enzymes (UbcD1). Sina also interacts with Tramtrack (Trk88) via PHYL. Tramtrack is a transcriptional repressor that blocks photoreceptor determination, while PHYL down-regulates the activity of Trk88. In turn, the activity of PHYL requires the activation of the Sevenless receptor tyrosine kinase, a process essential for R7 determination. It is thought that thus Sina targets Trk88 for degradation, therefore promoting the R7 pathway. Murine and human homologues of Sina have also been identified. The human homologue Siah-1 also binds E2 enzymes (UbcH5) and through a series of physical interactions, targets beta-catenin for ubiquitin degradation. Siah-1 expression is enhanced by p53, itself promoted by DNA damage. Thus this pathway links DNA damage to beta-catenin degradation. Sina proteins, therefore, physically interact with a variety of proteins. The N-terminal RING finger domain that binds ubiquitin conjugating enzymes is described in pfam00097, and does not form part of the alignment for this family. The remainder C-terminal part is involved in interactions with other proteins, and is included in this alignment. In addition to the Drosophila protein and mammalian homologues, whose similarity was noted previously, this family also includes putative homologues from Caenorhabditis elegans, Arabidopsis thaliana"

/db\_xref="CDD:pfam03145"

ORIGIN

Query Match 75.2%; Score 958; DB 9; Length 1540;

Best Local Similarity 100.0%; Pred. No. 4e-201;

Matches 958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

317 CAGAAATGAGCGGTGCTGCTACAGCATTAACCGTACTCGTCAAGTGTCCACCAT 376

113 CAGAAATGAGCGGTGCTGCTACAGCATTAACCGTACTCGTCAAGTGTCCACCAT 172

377 CCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436

173 CCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 232

437 TTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCATCTTCAATGTACAGTGGCC 496

233 TTGAGTGTCCAGTCTGCTTTGACTATGTGTATACCGCCCATCTTCAATGTACAGTGGCC 292

497 ATCTTGTGTAGCAACTGTGCGCCAAAGCTACATGTGTGCAACTTGCCTGGGCGCCCTT 556

293 ATCTTGTGTAGCAACTGTGCGCCAAAGCTACATGTGTGCAACTTGCCTGGGCGCCCTT 352

557 TGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTCTGTA 616

353 TGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTCTGTA 412

617 AATATGCGTCTTCTGCGATGTAATAAATCTCTGCGACACACAGAAAAAGCAGCATGAAG 676

413 AATATGCGTCTTCTGCGATGTAATAAATCTCTGCGACACACAGAAAAAGCAGCATGAAG 472

677 AGCTCTGTGAGTTAGGCTTATTCCTGTCGCGCTGCTTCTGTAATAAGTGGCAAG 736

473 AGCTCTGTGAGTTAGGCTTATTCCTGTCGCGCTGCTTCTGTAATAAGTGGCAAG 532

737 GCTCTCTGGATGCTGTAATGCGCCCATCTGATGTCATCAGCATAAATTCATCAACCTTAC 796

533 GCTCTCTGGATGCTGTAATGCGCCCATCTGATGTCATCAGCATAAATTCATCAACCTTAC 592

797 AGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTG 856

593 AGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTG 652

857 TGATGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916

653 TGATGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 712

917 ACGATGTCACACAGCAGTCTTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTG 976

713 ACGATGTCACACAGCAGTCTTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTG 772

977 AAAATTTTGTCTTACCGACTTGAGCTTAATGTCATAGGCGAGCTTGTAGTGGGAAGCGA 1036

773 AAAATTTTGTCTTACCGACTTGAGCTTAATGTCATAGGCGAGCTTGTAGTGGGAAGCGA 832

1037 CTCTCTCGATCTATTTCATGAAGGAATTCGCAACAGCCATTATGAATAGCGACTGTCTAGTCT 1096

833 CTCTCTCGATCTATTTCATGAAGGAATTCGCAACAGCCATTATGAATAGCGACTGTCTAGTCT 892

1097 TTGACACAGCAGTTCGACAGCTTTTTCGCAAAAATGGCAATTTAGGCATCAATGAACCTA 1156

893 TTGACACAGCAGTTCGACAGCTTTTTCGCAAAAATGGCAATTTAGGCATCAATGAACCTA 952

1157 TTTCCATGTTGTAATGGCAATCAACATTTTCTGGCCAGTGTATAAACTTCAGTTTC 1216

953 TTTCCATGTTGTAATGGCAATCAACATTTTCTGGCCAGTGTATAAACTTCAGTTTC 1012

1217 ACAGAAAAATAGGCGACCCCATCTGCTGCGCAACTTAACTTTTCGGTAGGTGGAGC 1274

1013 ACAGAAAAATAGGCGACCCCATCTGCTGCGCAACTTAACTTTTCGGTAGGTGGAGC 1070

RESULT 15

CO719559

LOCUS

Sequence 5493 from Patent WO02068579.

CO719559

ACCESSION

CO719559.1

GI:42280416

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS

TITLE

JOURNAL

Patent: WO 02068579-A 5493 06-SEP-2002;

CO719559 1874 bp DNA linear PAT 03-FEB-2004

Sequence 5493 from Patent WO02068579.

CO719559

ACCESSION

CO719559.1

GI:42280416

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS

TITLE

JOURNAL

Patent: WO 02068579-A 5493 06-SEP-2002;

FEATURES PE Corporation (NY) (US)  
Location/Qualifiers  
source 1..1874  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 75.2%; Score 958; DB 6; Length 1874;  
Best Local Similarity 100.0%; Pred. No. 3.9e-201;  
Matches 958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 CAGAAATGAGCCGTGAGACTGTACAGCACTTACCGGTACCTCGAAGTGTCCACCAT 376  
Db 1 CAGAAATGAGCCGTGAGACTGTACAGCACTTACCGGTACCTCGAAGTGTCCACCAT 60

QY 377 CCAGAGGGTGCCTGCCCTGACTGGCACAACCTGCAATCCAAATGACTTGGCGAGTCTTT 436  
Db 61 CCAGAGGGTGCCTGCCCTGACTGGCACAACCTGCAATCCAAATGACTTGGCGAGTCTTT 120

QY 437 TTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCAATCTTCAATGTCAAGTGGCC 496  
Db 121 TTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCAATCTTCAATGTCAAGTGGCC 180

QY 497 ATCTTGTGTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCAACTTGC CGGGGCCCTT 556  
Db 181 ATCTTGTGTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCAACTTGC CGGGGCCCTT 240

QY 557 TGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTCCCTGTA 616  
Db 241 TGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTCCCTGTA 300

QY 617 AATATGCGTCTTCTGGATGTGAATAAATCTCTGCCACACAGAAAAAGCAGACCATGAAG 676  
Db 301 AATATGCGTCTTCTGGATGTGAATAAATCTCTGCCACACAGAAAAAGCAGACCATGAAG 360

QY 677 AGCTCTGTGAGTTTGGCCCTTATTCCTGTGCGGCCCTGGTGCTTCTGTAATGGCAAG 736  
Db 361 AGCTCTGTGAGTTTGGCCCTTATTCCTGTGCGGCCCTGGTGCTTCTGTAATGGCAAG 420

QY 737 GCTCTCTGGATGCTGTAATGCCCCCATCTGATGATCAGCATAGTCCATTACAACTTAC 796  
Db 421 GCTCTCTGGATGCTGTAATGCCCCCATCTGATGATCAGCATAGTCCATTACAACTTAC 480

QY 797 AGGAGAGGATATAGTTTTTTCTTGCTACAGACATTAATCTTCTGGTGCTGTGACTGGG 856  
Db 481 AGGAGAGGATATAGTTTTTTCTTGCTACAGACATTAATCTTCTGGTGCTGTGACTGGG 540

QY 857 TGATGATGAGTCCCTGTTTTGGCTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAAAT 916  
Db 541 TGATGATGAGTCCCTGTTTTGGCTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAAAT 600

QY 917 ACGATGGTCAACAGCAGTTCTTGGCAATCGTACAGCTGATAGCAACAGCAAGCAAGCTG 976  
Db 601 ACGATGGTCAACAGCAGTTCTTGGCAATCGTACAGCTGATAGCAACAGCAAGCAAGCTG 660

QY 977 AAAATTTTCTTACGACTTGGCTTAAATGGTCAATAGGCGACGATTGACTTGGGAAGCGA 1036  
Db 661 AAAATTTTCTTACGACTTGGCTTAAATGGTCAATAGGCGACGATTGACTTGGGAAGCGA 720

QY 1037 CTCTCGATCTATTCAATGAAGAAATGCAACAGCCATTAATGAATAGCGACTGTCTAGTCT 1096  
Db 721 CTCTCGATCTATTCAATGAAGAAATGCAACAGCCATTAATGAATAGCGACTGTCTAGTCT 780

QY 1097 TTGACACCAAGCATTCGACAGCTTTTGGCAAAATGGCAATTTAGGCATCAATGTAACTA 1156  
Db 781 TTGACACCAAGCATTCGACAGCTTTTGGCAAAATGGCAATTTAGGCATCAATGTAACTA 840

QY 1157 TTTCCATGTGTTGAAATGGCAATCAAAATTTTCTGGCCAGTGTGTTTAAAACTTCAGTTTC 1216  
Db 841 TTTCCATGTGTTGAAATGGCAATCAAAATTTTCTGGCCAGTGTGTTTAAAACTTCAGTTTC 900

QY 1217 ACAGAAAAATAGGCAACCCATCTGTCTGCCAACCTTAAACTCTTCGGTAGGTGGAAGC 1274

Db 901 ACAGAAAAATAGGCAACCCCATCTGTCTGCCAACCTTAAACTCTTCGGTAGGTGGAAGC 958

Search completed: April 24, 2005, 21:36:35  
Job time : 5667 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 19:43:00 ; Search time 4597 Seconds  
(without alignments)  
10549.027 Million cell updates/sec

Title: US-10-679-246-1  
Perfect score: 1274  
Sequence: 1 tttctttagtggtttatgt.....ctcttcgtagtggaagc 1274

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hic:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gse1:  
9: gb\_gse2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1062	83.4	1221	7	U70056
2	958	75.2	1360	3	CR607323
3	958	75.2	1857	3	CR614925
4	894	70.2	1287	3	CR598465
5	849	66.6	849	9	AY402255
6	843.2	66.2	849	9	AY402256
7	832.4	65.3	921	5	EX355841
8	811	63.7	859	5	EX355841
9	749.8	58.9	849	9	AY402257
10	749.2	58.8	824	4	BG619351
11	748.8	58.8	779	1	AI936347
12	747.4	58.7	778	1	AI971258
13	733	57.5	872	5	EX448317
14	708.8	56.4	785	5	EX384656
15	708.8	55.6	852	5	EX463976
16	699.6	54.9	769	5	EX097203
17	693.6	54.4	714	5	BQ447177
18	686.2	53.9	890	7	CF551876
19	685	53.8	1576	3	CR593197
20	682.4	53.6	693	5	BU662334
21	680.2	53.4	1058	5	EX356371
22	656.6	51.5	848	5	BU614872
23	650.8	51.1	812	1	AU119916
24	645.8	50.7	649	2	BF446537

25	641.4	50.3	647	2	BE503207
26	640.6	50.3	800	4	BI836179
27	631.8	49.6	637	1	AI681134
28	627.4	49.2	647	2	BE502849
29	622	48.8	622	7	CV030563
30	621.8	48.8	715	5	EX459079
31	620.6	48.7	957	5	EX372197
32	620.4	48.7	953	6	CA980024
33	610.8	47.9	842	5	EX459114
34	606.6	47.6	853	4	BG436513
35	594.6	46.7	608	1	AI698102
36	590	46.3	735	4	BI560892
37	587.4	46.1	864	4	BG489183
38	586.2	46.0	716	5	BU260826
39	585.2	45.9	658	4	BG537816
40	582.8	45.7	902	4	BI763325
41	580.2	45.5	734	1	AU138740
42	575	45.1	1025	5	EX366403
43	571.2	44.8	767	4	BG435333
44	569.6	44.7	976	5	BU191771
45	562.4	44.1	894	2	BF573693

ALIGNMENTS

RESULT 1  
LOCUS U70056 1221 bp mRNA linear EST 27-OCT-1999  
DEFINITION U70056 Soares infant brain INIB Homo sapiens cDNA clone 25050, mRNA sequence.  
ACCESSION U70056  
VERSION U70056.1 GI:2731404  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Volorio, S., Simon, G., Repetto, M., Cucciarini, M., Banfi, S., Borgani, G., Ballabio, A. and Zollo, M.  
TITLE Sequencing analysis of forty-eight human image cDNA clones similar to drosophila mutant protein  
JOURNAL DNA Seq. 9 (5-6), 307-315 (1998)  
MEDLINE 99452388  
PubMed 10524757  
COMMENT Contact: Zollo, Massimo  
Telethon Institute of Genetics and Medicine  
Via Olgettina 58, Milan, MI 20132, Italy  
Email: zollo@tigem.it

FEATURES

source  
Location/Qualifiers  
1..1221  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="16q12.1-q13"  
/clone="25050"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares infant brain INIB"  
/note="Organ: whole brain; Vector: Lafmid BA; Site: 1: Not I; Site 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', RACTGGAGATTCGCGCCGACGAAATTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 83.4%; Score 1062; DB 7; Length 1221;

Best Local Similarity 100.0%; Pred. No. 2,5e-248; Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
2y	213	AAATTTTGTATGAAACTTTTAAAGAGCACTTATGGCATGTAAACATTTATTTATAAGTTAAGT	272
Db	1	AAATTTTGTATGAAACTTTTAAAGAGCACTTATGGCATGTAAACATTTATTTATAAGTTAAGT	60
2y	273	CATGGTTATTAATTTATTTTCTCTCGCTCCCTTATGTATTTATTTTCAGAAATGAGCCGTC	332
Db	61	CATGGTTATTAATTTATTTTCTCTCGCTCCCTTATGTATTTATTTTCAGAAATGAGCCGTC	120
2y	333	GACTGTACAGCATTAACCTACCGGTACCTCGAAAGTGTCCACCATCCAGAGGGTCCCTGC	392
Db	121	GACTGTACAGCATTAACCTACCGGTACCTCGAAAGTGTCCACCATCCAGAGGGTCCCTGC	180
2y	393	CCTGACTGGCAACAACCTGCAATCAACATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTG	452
Db	181	CCTGACTGGCAACAACCTGCAATCAACATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTG	240
2y	453	CTTTGACTATGTGTATACCGCCCATTTCAATGTCCAGAGTGCCCATCTGTGTTGTAGCAA	512
Db	241	CTTTGACTATGTGTATACCGCCCATTTCTCAATGTCCAGAGTGCCCATCTGTGTTGTAGCAA	300
2y	513	CTGTCCGCCAAAGCTCACATGTGTCCAACTTGC CGGGGCCCTTTGGATCCATTCGCAA	572
Db	301	CTGTCCGCCAAAGCTCACATGTGTGTCCAACTTGC CGGGGCCCTTTGGGATCCATTCGCAA	360
2y	573	CTTGGCTATGGAAGTGGCTAAATTCAGTACTTTTCCCTGTAATATGCGTCTTCTGG	632
Db	361	CTTGGCTATGGAAGTGGCTAAATTCAGTACTTTTCCCTGTAATATGCGTCTTCTGG	420
2y	633	ATGTGAAATAACTCTGTCCACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTT	692
Db	421	ATGTGAAATAACTCTGTCCACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTT	480
2y	693	GCCTTATTCCTGTGCGTGGCTTCCTGTAAATGGCAAGGCTCTCTGGATGCTGT	752
Db	481	GCCTTATTCCTGTGCGTGGCTTCCTGTAAATGGCAAGGCTCTCTGGATGCTGT	540
2y	753	AATGCCCATCTGATGCATCAGCAATGAAGTCCATTAACAACCTACAGGGAGAGATATAGT	812
Db	541	AATGCCCATCTGATGCATCAGCAATGAAGTCCATTAACAACCTACAGGGAGAGATATAGT	600
2y	813	TTTTTCTGTCTACAGCATTAATCTTCTCGGTGCTGTGTGACTGGGTGATGATCGAGTCTCTG	872
Db	601	TTTTTCTGTCTACAGCATTAATCTTCTCGGTGCTGTGTGACTGGGTGATGATCGAGTCTCTG	660
2y	873	TTTTTGGCTTTCACTTCATCTGTATGTCTTAGAGAAACAGGAAAAATACGATGTCAACAGCA	932
Db	661	TTTTTGGCTTTCACTTCATCTGTATGTCTTAGAGAAACAGGAAAAATACGATGTCAACAGCA	720
2y	933	GTTCTTTTCGCAATCGTACAGCTGATAGGAAACACGCAAGCAAGCTGAAAAATTTTGTCTTACCG	992
Db	721	GTTCTTTTCGCAATCGTACAGCTGATAGGAAACACGCAAGCAAGCTGAAAAATTTTGTCTTACCG	780
2y	993	ACTTGAAGTAAATGGTCTATAGGCGCATTTGACTTTGGGAGCGACTCTTCGATCTATTCTCA	1052
Db	781	ACTTGAAGTAAATGGTCTATAGGCGCATTTGACTTTGGGAGCGACTCTTCGATCTATTCTCA	840
2y	1053	TGAAGGAAATGGCAACAGCCATTAATGAATAGCGACTGTCTAGTCTTTTGACACCCAGCATTTGC	1112
Db	841	TGAAGGAAATGGCAACAGCCATTAATGAATAGCGACTGTCTAGTCTTTTGACACCCAGCATTTGC	900
2y	1113	ACAGCTTTTTTGCAGAAAAATGGCAATTTAGGCAATCAATGTAACTATTTTCCATGTGTGAAA	1172
Db	901	ACAGCTTTTTTGCAGAAAAATGGCAATTTAGGCAATCAATGTAACTATTTTCCATGTGTGAAA	960
2y	1173	TGGCAATCAACATTTTCTGCGCAGTGTTTTAAAACTTCAGTTCACAGAAAAATAAGGCAC	1232
Db	961	TGGCAATCAACATTTTCTGCGCAGTGTTTTAAAACTTCAGTTCACAGAAAAATAAGGCAC	1020
2y	1233	CCATCTGTCTGCCAACCTATAAACTTTTTCGGTAGTGGAGC	1274

Db 1021 CCAATCTGTCGCCAACCTAAACTCTTTCGGTAGGTGGAAGC 1062

RESULT 2  
CR607323 1360 bp mRNA linear HTC 21-JUL-2004  
LOCUS full-length cDNA clone CSODC023YH14 of Neuroblastoma Cot  
DEFINITION 25-normalized of Homo sapiens (human).  
ACCESSION CR607323.1 GI:50488130  
VERSION HTCCNSLT CDNA.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1360)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ invitrogen Corporation 1600  
Faraday Avenue  
2 (bases 1 to 1360)  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequences :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five primer  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of invitrogen.  
FEATURES  
source location/Qualifiers  
1..1360  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODC023YH14"  
/tissue\_type="Neuroblastoma Cot 25-normalized"  
/plasmid="pCMVSPORT\_6"

ORIGIN  
Query Match 75.2%; Score 958; DB 3; Length 1360;  
Best Local Similarity 100.0%; Pred. No. 6.5e-223;  
Matches 958; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 317 CAGAAATGAGCGTCAGACTGCTACAGCATTTACTACCGGTACCTCGAAGTCTCCACCAT 376  
|||||  
Db 1 CAGAAATGAGCGTCAGACTGCTACAGCATTTACTACCGGTACCTCGAAGTCTCCACCAT 60  
|||||

QY 377 CCAGAGGTGCTGCCCTGACTGGCACAACCTGCATCCAAACATGACTTTGGCGAGTCTTT 436  
|||||

Db 61 CCAGAGGTGCTGCCCTGACTGGCACAACCTGCATCCAAACATGACTTTGGCGAGTCTTT 120  
|||||

QY 437 TTGAGTGTCCAGTCTGCTTTTGACTATGTTTACCGCCCATTTCTTCAATGTCAAGTGGCC 496  
|||||

Db 121 TTGAGTGTCCAGTCTGCTTTTGACTATGTTTACCGCCCATTTCTTCAATGTCAAGTGGCC 180  
|||||

QY 497 ATCTTGTTTGTAGCAACTGCTGCCCAAAAGCTCACATGTTTGTCCAACTTCGCGGGGCCCTT 556  
|||||

Db 181 ATCTTGTTTGTAGCAACTGCTGCCCAAAAGCTCACATGTTTGTCCAACTTCGCGGGGCCCTT 240  
|||||

QY 557 TGGGATCCATTCCGCAACTTTGGCTATGGAGAAAGTGGCTAATTCAAGTACTTTTCCCTGTGA 616  
|||||

Db 241 TGGGATCCATTCCGCAACTTTGGCTATGGAGAAAGTGGCTAATTCAAGTACTTTTCCCTGTGA 300  
|||||

QY 617 AATATGCGTCTTCGATGTGAATAAATCTCTGCCACACAGAGAAAGCAGACCATGAAG 676  
|||||

Db 301 AATATGCGTCTTCGATGTGAATAAATCTCTGCCACACAGAGAAAGCAGACCATGAAG 360  
|||||

QY 677 AGCTCTGTGAGTTTAGCGCCTTAATTCCTGTGCGCCTGTGTTGCTTCTGTAAATGGCAAG 736  
|||||





QY	801	AGAGGATATAGTTTTTTCTTGCTACAGCAATTAATCTTCTGGTGGCTGTGACATGGGTGAT	860
Db	421	AGAGGATATAGTTTTTTCTTGCTACAGCAATTAATCTTCTGGTGGCTGTGACATGGGTGAT	480
QY	861	GATGCAGTCCTGTTTTGGCTTTTACATTCATGTTAGTCTTTAGAGAAAACAGGAAAAATACGA	920
Db	481	GATGCAGTCCTGTTTTGGCTTTTACATTCATGTTAGTCTTTAGAGAAAACAGGAAAAATACGA	540
QY	921	TGCTCACACGACGCTTCTTCGCAATCGTACAGCTGTATAGGAACACGCAAGCAAGCTGAAAA	980
Db	541	TGCTCACACGACGCTTCTTCGCAATCGTACAGCTGTATAGGAACACGCAAGCAAGCTGAAAA	600
QY	981	TTTTTGCTTACCGACTTTCAGCTAAATGGTCTATAGGCAAGCATTCAGCTTGGGAAGCACTCC	1040
Db	601	TTTTTGCTTACCGACTTTCAGCTAAATGGTCTATAGGCAAGCATTCAGCTTGGGAAGCACTCC	660
QY	1041	TCCATCTATTTCATGAAGGAATTCGAACAGCCATTAATGAATAGGACATGTCTAGTCTTTTGA	1100
Db	661	TCCATCTATTTCATGAAGGAATTCGAACAGCCATTAATGAATAGGACATGTCTAGTCTTTTGA	720
QY	1101	CACGACATTCGACAGCTTTTTTTCGAGAAAATGGCAATTTAGGATCAATGTAAGTCAATTTTC	1160
Db	721	CACGACATTCGACAGCTTTTTTTCGAGAAAATGGCAATTTAGGATCAATGTAAGTCAATTTTC	780
QY	1161	CATCTGTTGAAATGGCAATCAAAACATTTCTGGCCAGTGTTTAAACCTTCAGTTTTCACAG	1220
Db	781	CATCTGTTGAAATGGCAATCAAAACATTTCTGGCCAGTGTTTAAACCTTCAGTTTTCACAG	840
QY	1221	AAAATAAGGCACCCATCTGTCTGCCAACCTTAAACCTTTTCGGTAGTGGAAGC	1274
Db	841	AAAATAAGGCACCCATCTGTCTGCCAACCTTAAACCTTTTCGGTAGTGGAAGC	894
RESULT 5	AY402255	849 bp DNA linear	GSS 15-DEC-2003
LOCUS	DEFINITION	Homo sapiens SIAH1 gene, VIRTUAL TRANSCRIPT, partial sequence,	
ACCESSION	GENOMIC SURVEY SEQUENCE		
VERSION	AY402255		
KEYWORDS	AY402255.1 GI:39758241		
SOURCE	GSS		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
	1 (bases 1 to 849)		
	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 849)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..849		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	<1..>849		
gene	/gene="SIAH1"		
	/locus_tag="HCW1163"		

```

RESULT 4
CR598465      1287 bp      mRNA      linear      HTC 21-JUL-2004
full-length cDNA clone CS0D1004YF12 of Placenta Cot 25-normalized
LOCUS       1287 bp      mRNA      linear      HTC 21-JUL-2004
DEFINITION   full-length cDNA clone CS0D1004YF12 of Placenta Cot 25-normalized
ACCESSION   CR598465
VERSION     CR598465.1 GI:50479272
KEYWORDS    HTC; CNSLT_CDNA.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1287)
            Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
            Full-length cDNA libraries and normalization
            Unpublished
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue
            2 (bases 1 to 1287)
            Genoscope.
            Direct Submission
            Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the PCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            Location/Qualifiers
            1..1287
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="CS0D1004YF12"
               /tissue_type="Placenta Cot 25-normalized"
               /plasmid="PCMVSPORT_6"

FEATURES             Source
ORIGIN
Query Match       70.2%; Score 894; DB 3; Length 1287;
Best Local Similarity 100.0%; Pred. No. 2.7e-207; Indels 0; Gaps 0;
Matches 894; Conservative 0; Mismatches 0;

381  GAGGGTGCCTGCCCTGACTGGCAAACTGATGTTGTCCAACTTCCCGGGGCCCTTTGGG 440
    1  GAGGGTGCCTGCCCTGACTGGCAAACTGATGTTGTCCAACTTCCCGGGGCCCTTTGGG 60
    441  GTGTCCAGTCTGCTTTGACTATGTTTACCGCCCATCTTCAATGTCAGAGTGCCCATCT 500
    61  GTGTCCAGTCTGCTTTGACTATGTTTACCGCCCATCTTCAATGTCAGAGTGCCCATCT 120
    501  TGTTTGTAGCAACTGTGCCCAAGCTCACATGTTGTCCAACTTCCCGGGGCCCTTTGGG 560
    121  TGTTTGTAGCAACTGTGCCCAAGCTCACATGTTGTCCAACTTCCCGGGGCCCTTTGGG 180
    561  ATCCATTGCGAACTGTGGCTATGGAGAAAGTGCTTAATTCAGTACTTTTCCCTCTGAATA 620
    181  ATCCATTGCGAACTGTGGCTATGGAGAAAGTGCTTAATTCAGTACTTTTCCCTCTGAATA 240
    621  TGCGCTCTTCGGATGTGAATAAATCTCTGCCACACACAGAAAAGCAGACCATGAAGAGCT 680
    241  TGCCTCTTCTGGATGTGAATAAATCTCTGCCACACACAGAAAAGCAGACCATGAAGAGCT 300
    681  CTGTGAGTTTAGGCCTTAATTCCTGTCCGTGCGCTTCTTCGTAAATGGCAGGCTC 740
    301  CTGTGAGTTTAGGCCTTAATTCCTGTCCGTGCGCTTCTTCGTAAATGGCAGGCTC 360
    741  TCTGGATGCTGAATGCCCATCTGATCATCAGCATAAGTCCATTACAAACCTACAGGG 800
    361  TCTGGATGCTGAATGCCCATCTGATCATCAGCATAAGTCCATTACAAACCTACAGGG 420

```



ORIGIN

Query Match 66.6%; Score 849; DB 9; Length 849;  
Best Local Similarity 100.0%; Pred. No. 2.5e-196;  
Matches 849; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 ATGAGCGCTGACTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACATCCCGAG 381  
DB 1 ATGAGCGCGTACAGCTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACATCCCGAG 60

QY 382 AGGGTCCCTGCGCTGACTGGCACTGATCGATCAACATGATGCTGGCGAGTCTTTTGGAG 441  
DB 61 AGGGTCCCTGCGCTGACTGGCACTGATCGATCAACATGATGCTGGCGAGTCTTTTGGAG 120

QY 442 TGTCCAGTCTGCTTTGACTATGTTTACCGCCCATTTCTTCAATGTCCAGAGTGCCTATCTT 501  
DB 121 TGTCCAGTCTGCTTTGACTATGTTTACCGCCCATTTCTTCAATGTCCAGAGTGCCTATCTT 180

QY 502 GTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCGAATGTCGCGGGCCCTTTGGGA 561  
DB 181 GTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCGAATGTCGCGGGCCCTTTGGGA 240

QY 562 TCCATTCGCAACTGTGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTCTGTAATAT 621  
DB 241 TCCATTCGCAACTGTGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTCTGTAATAT 300

QY 622 GCGTCTTCTGGATGTAATAACTCTGCCACACACAGAAAAGCAGACCATGAAGAGCTC 681  
DB 301 GCGTCTTCTGGATGTAATAACTCTGCCACACACAGAAAAGCAGACCATGAAGAGCTC 360

QY 682 TGTGAGTTTGGCCCTTATTCCTGTCGCTGCGCTGCTTCTCTGTAATGCGAGGCTCT 741  
DB 361 TGTGAGTTTGGCCCTTATTCCTGTCGCTGCGCTGCTTCTCTGTAATGCGAGGCTCT 420

QY 742 CTGGATGCTGTAATGCCCATCTGATGCATCAGCAATGCTCAATACCAACCTTACAGGCA 801  
DB 421 CTGGATGCTGTAATGCCCATCTGATGCATCAGCAATGCTCAATACCAACCTTACAGGCA 480

QY 802 GAGGATATAGTTTTCTTGCTACAGACATTAATCTTCTGCTGCTGTTGACTGGGTGATG 861  
DB 481 GAGGATATAGTTTTCTTGCTACAGACATTAATCTTCTGCTGCTGTTGACTGGGTGATG 540

QY 862 ATGCACTCTGTTTGGCTTTTCACTTCATGTTAGTCTTACAGAAACAGGAAAAATACGAT 921  
DB 541 ATGCACTCTGTTTGGCTTTTCACTTCATGTTAGTCTTACAGAAACAGGAAAAATACGAT 600

QY 922 GGTCAACAGCTTCTTCCCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAATAT 981  
DB 601 GGTCAACAGCTTCTTCCCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAATAT 660

QY 982 TTTGCTTACCGACTTGAGCTAAATGGTCTATAGGCGACGATTTGACTTTGGGAGCGACTCT 1041  
DB 661 TTTGCTTACCGACTTGAGCTAAATGGTCTATAGGCGACGATTTGACTTTGGGAGCGACTCT 720

QY 1042 CGATCTATTCATGAGGAATTCGACAGCCATTAATGATAGGCACTGCTAGTCTTTGAC 1101  
DB 721 CGATCTATTCATGAGGAATTCGACAGCCATTAATGATAGGCACTGCTAGTCTTTGAC 780

QY 1102 ACCAGCATTCACAGCTTTTTCGAGAAATGCAATTTAGGCGATCAATGTAATTTTCC 1161  
DB 781 ACCAGCATTCACAGCTTTTTCGAGAAATGCAATTTAGGCGATCAATGTAATTTTCC 840

QY 1162 ATGTGTTGA 1170  
DB 841 ATGTGTTGA 849

RESULT 6

AY402256  
LOCUS Pan troglodytes STAH1 gene, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION AY402256

AY402256.1 GI:39758242  
GSS.  
Pan troglodytes (chimpanzee)  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
REFERENCE  
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE  
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source  
1..849  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>849  
/gene="STAH1"  
/locus\_tag="HCM1163"  
ORIGIN  
Query Match 66.2%; Score 843.2; DB 9; Length 849;  
Best Local Similarity 99.5%; Pred. No. 6.5e-195;  
Matches 845; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 322 ATGAGCGCTGACTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACATCCCGAG 381  
DB 1 ATGAGCGCGTACAGCTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACATCCCGAG 60

QY 382 AGGGTCCCTGCGCTGACTGGCACTGATCGATCAACATGATGCTGGCGAGTCTTTTGGAG 441  
DB 61 AGGGTCCCTGCGCTGACTGGCACTGATCGATCAACATGATGCTGGCGAGTCTTTTGGAG 120

QY 442 TGTCCAGTCTGCTTTGACTATGTTTACCGCCCATTTCTTCAATGTCCAGAGTGCCTATCTT 501  
DB 121 TGTCCAGTCTGCTTTGACTATGTTTACCGCCCATTTCTTCAATGTCCAGAGTGCCTATCTT 180

QY 502 GTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCGAATGTCGCGGGCCCTTTGGGA 561  
DB 181 GTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCGAATGTCGCGGGCCCTTTGGGA 240

QY 562 TCCATTCGCAACTGTGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTCTGTAATAT 621  
DB 241 TCCATTCGCAACTGTGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTCTGTAATAT 300

QY 622 GGGTCTTCTGGATGTAATAACTCTGCCACACACAGAAAAGCAGACCATGAAGAGCTC 681  
DB 301 GGGTCTTCTGGATGTAATAACTCTGCCACACACAGAAAAGCAGACCATGAAGAGCTC 360

QY 682 TGTGAGTTTAGGCGCTTATTCCTGTCGCTGCGCTGCTTCTCTGTAATGCGAGGCTCT 741  
DB 361 TGTGAGTTTAGGCGCTTATTCCTGTCGCTGCGCTGCTTCTCTGTAATGCGAGGCTCT 420

QY 742 CTGGATGCTGTAATGCCCATCTGATGCATCAGCAATGCTCAATACCAACCTTACAGGCA 801  
DB 421 CTGGATGCTGTAATGCCCATCTGATGCATCAGCAATGCTCAATACCAACCTTACAGGCA 480

QY 802 GAGGATATAGTTTTCTTGCTACAGACATTAATCTTCTGCTGCTGTTGACTGGGTGATG 861  
DB 481 GAGGATATAGTTTTCTTGCTACAGACATTAATCTTCTGCTGCTGTTGACTGGGTGATG 540

QY 862 ATGCACTCTGTTTGGCTTTTCACTTCATGTTAGTCTTACAGAAACAGGAAAAATACGAT 921  
DB 541 ATGCACTCTGTTTGGCTTTTCACTTCATGTTAGTCTTACAGAAACAGGAAAAATACGAT 600

QY 922 GGTCAACAGCTTCTTCCCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAATAT 981  
DB 601 GGTCAACAGCTTCTTCCCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAATAT 660

QY 982 TTTGCTTACCGACTTGAGCTAAATGGTCTATAGGCGACGATTTGACTTTGGGAGCGACTCT 1041  
DB 661 TTTGCTTACCGACTTGAGCTAAATGGTCTATAGGCGACGATTTGACTTTGGGAGCGACTCT 720

QY 1042 CGATCTATTCATGAGGAATTCGACAGCCATTAATGATAGGCACTGCTAGTCTTTGAC 1101  
DB 721 CGATCTATTCATGAGGAATTCGACAGCCATTAATGATAGGCACTGCTAGTCTTTGAC 780

QY 1102 ACCAGCATTCACAGCTTTTTCGAGAAATGCAATTTAGGCGATCAATGTAATTTTCC 1161  
DB 781 ACCAGCATTCACAGCTTTTTCGAGAAATGCAATTTAGGCGATCAATGTAATTTTCC 840

QY 1162 ATGTGTTGA 1170  
DB 841 ATGTGTTGA 849

Db 481 GAGGATATAGTTTCTTCTGCTACAGACATTAACTTCTCGTGTGTTGACTGGGTGATG 540  
QY 862 ATGCAGTCTCTGTTTGGCTTTCACCTTCATGTTAGTCTTGTAGAGAAACAGGAAAAATACGAT 921  
Db 541 ATGCAGTCTCTGTTTGGCTTTCACCTTCATGTTAGTCTTGTAGAGAAACAGGAAAAATACGAT 600  
QY 922 GGTCAACAGCAGTCTTCTGCAATCTGACAGCTGATAGGAAACAGCAAGCAAGCTGAAAAAT 981  
Db 601 GGTCAACAGCAGTCTTCTGCAATCTGACAGCTGATAGGAAACAGCAAGCAAGCTGAAAAAT 660  
QY 982 TTTTGTACCGACTTGTAGCTAAATGGTATAGCGGACGATGACCTTGGGAGCGACTCCT 1041  
Db 661 TTTGCTTACCGACTTGTAGCTAAATGGTATAGCGGACGATGACCTTGGGAGCGACTCCT 720  
QY 1042 CGATCTATTTCATGAAGGAATTCACAGCATTATGAATAGGAGCTGTCTAGTCTTTGAC 1101  
Db 721 CGATCTATTTCATGAAGGAATTCACAGCATTATGAATAGGAGCTGTCTAGTCTTTGAC 780  
QY 1102 ACAGCAATGCAAGCTTTTTCAGAAATGCAATTTAGGCATCAATGTAATTAATTTCC 1161  
Db 781 ACCAGCAATGCAAGCTTTTTCAGAAATGCAATTTAGGCATCAATGTAATTAATTTCC 840  
QY 1162 ATGTGTTGA 1170  
Db 841 ATGTGTTGA 849

RESULT 7  
BX355841  
LOCUS  
DEFINITION  
BX355841 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1004YF12 5-PRIME, mRNA sequence.  
ACCESSION  
BX355841  
VERSION  
BX355841.2 GI:46288391  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 921)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization.  
Unpublished (2001)  
On May 5, 2003 this sequence version replaced gi:30369970.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
8104.r

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0D1004YF12&c=8104.r.  
FEATURES  
Location/Qualifiers  
1..921  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1004YF12"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 65.3%; Score 832.4; DB 5; Length 921;  
Best Local Similarity 95.8%; Pred. No. 2.9e-192;  
Matches 854; Conservative 13; Mismatches 20; Indels 4; Gaps 2;

QY 381 GAGGTGCTCCCTGACTGGCACTGCATCAACAATGACTTGGCGAGTCTTTTGA 440  
Db 1 GAGGTGCTCCCTGACTGGCACTGCATCAACAATGACTTGGCGAGTCTTTTGA 60  
QY 441 GTGTCCAGTCTGCTTTGACTATGTTTACCGCCCAATCTTCAATGTGAGAGTGGCCATCT 500  
Db 61 GTGTCCAGTCTGCTTTGACTATGTTTACCGCCCAATCTTCAATGTGAGAGTGGCCATCT 120  
QY 501 TGTGTTAGCACTGTGCGCCCAAGCTCACAATGTTGCTCAATGTTGGGCGCCCTTTGGG 560  
Db 121 TGTGTTAGCACTGTGCGCCCAAGCTCACAATGTTGCTCAATGTTGGGCGCCCTTTGGG 180  
QY 561 ATCCATTGCGAACTTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTAAATA 620  
Db 181 ATCCATTGCGAACTTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTAAATA 240  
QY 621 TGCCTCTCTGAGTGTGAATAAATCTGCGCACACACAGAAAAAGCAGACCAATGAGAGCT 680  
Db 241 TGCCTCTCTGAGTGTGAATAAATCTGCGCACACACAGAAAAAGCAGACCAATGAGAGCT 300  
QY 681 CTGTGAGTTTGGGCTTATTCCTGCTGCGCCCTGCTTCTCTGTAATGCGCAAGGCTC 740  
Db 301 CTGTGAGTTTGGGCTTATTCCTGCTGCGCCCTGCTTCTCTGTAATGCGCAAGGCTC 360  
QY 741 TCTGATGCTCTAATGCCCCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 800  
Db 361 TCTGATGCTCTAATGCCCCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 801 AGAGATATATGTTTCTGCTACAGCAATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 860  
Db 421 AGAGATATATGTTTCTGCTACAGCAATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 861 GATGAGCTCTGTTTGGCTTTCATCTCATGTTAGTCTTTAGAGAAACAGGAAAAATACGA 920  
Db 481 GATGAGCTCTGTTTGGCTTTCATCTCATGTTAGTCTTTAGAGAAACAGGAAAAATACGA 540  
QY 921 TGTGTCACAGCAGTCTTTCGCAATCGTACAGTGTATGAGAAACAGCAAGCAAGCTGAAAA 980  
Db 541 TGTGTCACAGCAGTCTTTCGCAATCGTACAGTGTATGAGAAACAGCAAGCAAGCTGAAAA 600  
QY 981 TTTTCTTACCGACTTGTAGCTAAATGGTCTATGAGCGAGCTTGTAGCTTGGGAGCGACTCC 1040  
Db 601 TTTTCTTACCGACTTGTAGCTAAATGGTCTATGAGCGAGCTTGTAGCTTGGGAGCGACTCC 660  
QY 1041 TCGATCTATTTCATGAAGGAATTCGCAAGCCCAATTAATGATGAGCTGTCTAGTCTTTGA 1100  
Db 661 TCGATCTATTTCATGAAGGAATTCGCAAGCCCAATTAATGATGAGCTGTCTAGTCTTTGA 720  
QY 1101 CACCGCAATGCAAGCTTTTTCGAGAAATGCGCAATTTAGGCATCAATGTAATTAATTTTC 1160  
Db 721 CACCGCAATGCAAGCTTTTTCGAGAAATGCGCAATTTAGGCATCAATGTAATTAATTTTC 780  
QY 1161 CATGTTTGAATGCAATCAACAATTTCTGCGCAGTGTTTTAAACCTTCAGTTTCACAG 1220  
Db 781 CATGTTTGAATGCAATCAACAATTTCTGCGCAGTGTTTTAAACCTTCAGTTTCACAG 840  
QY 1221 AAAATAAGCACCCTCTGCTGCGCAACCTAAACCTTTTCGTTAGGTGGA 1271  
Db 841 AAAATA--GGCACCCTCTGCTGCGCA--CCTAAACCTCTCTCGTAGGTGGA 887

RESULT 8  
BX366404  
LOCUS  
DEFINITION  
BX366404 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
cDNA clone CS0D023YH14 5-PRIME, mRNA sequence.  
ACCESSION  
BX366404  
VERSION  
BX366404.2 GI:46572835  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
AUTHORS      Mammalia: Eutheria: Primates; Catarrhini; Hominidae; Homo.
TITLE        1 (bases 1 to 859)
JOURNAL      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
COMMENT      Full-length cDNA libraries and normalization
              Unpublished (2001)
              On May 8, 2003 this sequence version replaced gi:30447532.
              Contact: Genoscope
              Genoscope - Centre National de Sequencage
              2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
              end enriched, double-strand cDNA was digested with Not I and cloned
              into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
              was normalized. Library was constructed by Life Technologies, a
              division of Invitrogen. This sequence belongs to sequence cluster
              8104.r
              For more information about this cluster, see
              http://www.genoscope.cns.fr/cdna?b=CSOAC023DD07Q28c=8104.r.

FEATURES
    source
        1..859
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CSOAC023YH14"
            /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was
            normalized."

ORIGIN
Query Match      63.7%; Score 811; DB 5; Length 859;
Best Local Similarity 98.5%; Pred. No. 4.8e-187;
Matches 844; Conservative 5; Mismatches 5; Indels 3; Gaps 3;

QY 317 CAGAATGAGCGGTGACAGTGTACAGATTAACCTACCGGTACCTCGAAGTGTCCACCAT 376
DB 1 CAGAATGAGCGGTGACAGTGTACAGATTAACCTACCGGTACCTCGAAGTGTCCACCAT 60
QY 377 CCAGAGGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
DB 61 CCAGAGGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 437 TTGAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
DB 121 TTGAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 497 ATCTTGTGTTGAGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
DB 181 ATCTTGTGTTGAGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 557 TGGGATCCATTCGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
DB 241 TGGGATCCATTCGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 617 AATATGCGTCTTCTGGATGTGAATAAATCTCTGCCACACACAGAAAGACAGACCATGAAG 676
DB 301 AATATGCGTCTTCTGGATGTGAATAAATCTCTGCCACACACAGAAAGACAGACCATGAAG 360
QY 677 AGCTCTGTGAGTTTGGCTTAAATCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
DB 361 AGCTCTGTGAGTTTGGCTTAAATCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 737 GCTCTCTGATGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
DB 421 GCTCTCTGATGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 797 AGGAGAGGATATAGTTTCTTGTGTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCT 856
DB 481 AGGAGAGGATATAGTTTCTTGTGTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCT 540
QY 857 TGATGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916

```

```

Db 541 TGATGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 917 ACGATGGTCCACGACGAGTCTTTCGCAATCGTACAGCTGTATAGGAAACGCAAGCAAGCTG 976
Db 601 ACGATGGTCCACGACGAGTCTTTCGCAATCGTACAGCTGTATAGGAAACGCAAGCAAGCTG 660
QY 977 AAATTTTGGCTTACGACGCTGAGCT-AAATGTCATAGCGGAGTGAATTTGGGAAAGC 1034
Db 661 AAATTTTGGCTTACGACGCTGAGCTTAAATGTCATAGCGGAGTGAATTTGGGAAAGC 720
QY 1035 GACTCTCTGATCTTATCATGAGGAAATGCAACAGCCATTAATGATAGGACGCTGCTGCTGCT 1094
Db 721 GACTCTCTGATCTTATCATGAGGAAATGCAACAGCCATTAATGATAGGACGCTGCTGCTGCT 780
QY 1095 CTTTGACACGACGATTCACAGCTTTTTCGCAAAATGCAATTTAGGCAATCAATGTAAC 1154
Db 781 CTTTGACACGACGATTCACAGCTTTTTCGCAAAATGCAATTTAGGCAATCAATGTAAC 839
QY 1155 TATTTCCATGTTGAA 1171
Db 840 TATTTCCATGTTGAA 856

RESULT 9
AY402257
LOCUS
DEFINITION Mus musculus Slah1 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY402257
VERSION AY402257.1
KEYWORDS GI:39758243
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 849)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 849)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submision
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
    source
        1..849
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            <1..>849
            /gene="Slah1"
            /locus_tag="HGM1163"
    ORIGIN
        Query Match 58.9%; Score 749.8; DB 9; Length 849;
        Best Local Similarity 92.7%; Pred. No. 4.3e-172;
        Matches 787; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 322 ATGACCGCTCAGACTGTACAGCATTACCTACCGTACTCGAGTGTCCACCATCCAG 381
Db 1 ATGACCGCTCAGACTGTACAGCATTACCTACCGTACTCGAGTGTCCACCATCCAG 60

```



```
Db 541 GCAATCGTACAGCTGATAGGAACACCAAGCAAGCTTGAATAATTTCTTACCAGCTTGAG 600
Qy 1000 CTAATGCGTCATAGGCGCAGATTGACTTGGGAAGCGACTCCTCGATCTATTATCATGAAGCA 1059
Db 601 CTAATGCGTCATAGGCGCAGATTGACTTGGGAAGCGACTCCTCGATCTATTATCATGAAGCA 660
Qy 1060 ATTGCAACGCCATT-ATGAATAGCGACTGT-CTAGCTTTTGACACGAGATTGCAACAGC 1117
Db 661 ATTGCAACGCCATTATGAATAGCGACTGTCTAGCTTTTGACACGAGATTGCAACAGC 720
Qy 1118 TTTTTCGAGAAATGCAATGATAGGATCAATCTAACTATTCCATGTCTGGAATGGCA 1177
Db 721 -TTTTCGAGAAATGCAATGATAGGCTCCTAGCTAACTATATCCATGTGGTGAATGGC 779
Qy 1178 ATCAAAACATTTCTGCGCCAGTGTTTAAACCTTCAGTTTCCACAGAAA 1223
Db 780 AATCAAAACATATCTGCCAGTGGTTAAAA-TTCAGTTTCCACAGAAA 824

RESULT 11
AI936347
LOCUS
DEFINITION
  wo81b11.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2461725 3'
  similar to TR:043269 O43269 HSI4H1.1; mRNA sequence.
ACCESSION
  AI936347
VERSION
  AI936347.1 GI:5675217
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 779)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 1411 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 444.
  Location/Qualifiers
    1..779
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:2461725"
      /lab_host="DH10B"
      /clone_lib="NCI CGAP Kid11"
      /note="Organ: Kidney; Vector: pTT3D-Pac (Pharmacia) with
      a modified polylinker; Site: Not 1; Site 2: Eco RI;
      Plasmid DNA from the normalized library NCI CGAP Kid1 was
      prepared, and 88 circles were made in vitro. Following HAP
      purification, this DNA was used as tracer in a subtractive
      hybridization reaction. The driver was PCR-amplified cDNAs
      from a pool of 5,000 clones made from the same library
      (cloneIDs 1322376-1323911, 1456007-1456775, and
      1500552-1502855). Subtraction by Bento Soares and M.
      Fatima Bonaldo."
```

## ORIGIN

```
Query Match 58.8%; Score 748.8; DB 1; Length 779;
Best Local Similarity 97.4%; Pred. No. 7.4e-172;
Matches 759; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
```

```
Qy 203 TTTGCTCAGTAATAATTTTGTATGAAACTTTTAAAGGACTTATGGCATGTAAACATTATTTA 262
Db 1 TTTGCTCAGTAATAATTTTGTATGAAACTTTTAAAGGACTTATGGCATGTAAACATTATTTA 60
Qy 263 TAAAGTAAAGTCATGGTTATAATTTATTTTCTCCTGCTCCTTATGTATTTATTTTCAGAAA 322
Db 61 TAAAGTAAAGTCATGGTTATAATTTATTTTCTCCTGCTCCTTATGTATTTATTTTCAGAAA 120
Qy 323 TGAGCGCTCAGACTGCTACAGCAATTACCTACCGGTACCTCGAAGTGTCCACCATCCAGA 382
Db 121 TGAGCGCTCAGACTGCTACAGCAATTACCTACCGGTACCTCGAAGTGTCCACCATCCAGA 180
Qy 383 GGGTGCCTGCTCCTGACTGGCAACAATGCAATCCAAATGACTTTGGGAGTCTTTTTGAGT 442
Db 181 GGGTGCCTGCTCCTGACTGGCAACAATGCAATCCAAATGACTTTGGGAGTCTTTTTGAGT 240
Qy 443 GTCCAGTCTGCTTTGACTATGTTTACCGCCCACTTCTCAATGTCAAGTGTGGCCATCTTG 502
Db 241 GTCCAGTCTGCTTTGACTATGTTTACCGCCCACTTCTCAATGTCAAGTGTGGCCATCTTG 300
Qy 503 TTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCAACTTGCCTGGGGCCCTTTGGGAT 562
Db 301 TTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCAACTTGCCTGGGGCCCTTTGGGAT 360
Qy 563 CCATTGCGCAACTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTAAATATG 622
Db 361 CCATTGCGCAACTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTAAATATG 420
Qy 623 CCTCTCTCGATGTGAAATAAATCTGCCACACACAGAAAAGCAGACCATGAGAGCTCT 682
Db 421 CCTCTCTCGATGTGAAATAAATCTGCCACACACAGAAAAGCAGACCATGAGAGCTCT 480
Qy 683 GTGAGTTAGGCTTATTTCTGCTGCGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
Db 481 GTGAGTTAGGCTTATTTCTGCTGCGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 743 TCGATGCTGTAATGCGCCCACTGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 802
Db 541 TCGATGCTGTAATGCGCCCACTGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 600
Qy 803 AGGATATAGTTTCTTCTGCTACAGCAATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 862
Db 601 AGGATATAGTTTCTTCTGCTACAGCAATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 863 TGCAGTCTGTTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 922
Db 661 TGCAGTCTGTTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
Qy 923 GTCACGAGCAGTTCTTTCGCAATCGTACAGTGTATGAAACACGCAAGCAAGCTGAAAT 981
Db 721 GTCACGAGCAGTTCTTTCGCAATCGTACAGTGTATGAAACACGCTAGCAAGCAAGCTGAAAT 779
```

## RESULT 12

```
AI971258
LOCUS
DEFINITION
  wr27b08.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:248887 3'
  similar to TR:043269 O43269 HSI4H1.1; mRNA sequence.
ACCESSION
  AI971258
VERSION
  AI971258.1 GI:5768084
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 778)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
```

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbnp/image/image.html](http://www-bio.llnl.gov/bbnp/image/image.html)  
Insert Length: 826 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence set: 455.

FEATURES  
source

```

Location/Qualifiers
1. 778
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:248887"
  /sex="male"
  /sex_stage="adult"
  /lab_host="DH10B"
  /clone_lib="NCI_CGAP_P228"
  /note="organ: prostate; vector: modified polylinker with a normalized library NCI_CGAP_P228. The driver was produced by PCR amplification of the prostate cancer cell line PC9. The circles were made in vitro. This DNA was used as template for PCR amplification. The driver was produced from 5,000 clones made from 985608-986759, 1101192-1101193, 1101194-1101195, 1101196-1101197, 1101198-1101199, 1101200-1101201, 1101202-1101203, 1101204-1101205, 1101206-1101207, 1101208-1101209, 1101210-1101211, 1101212-1101213, 1101214-1101215, 1101216-1101217, 1101218-1101219, 1101220-1101221, 1101222-1101223, 1101224-1101225, 1101226-1101227, 1101228-1101229, 1101230-1101231, 1101232-1101233, 1101234-1101235, 1101236-1101237, 1101238-1101239, 1101240-1101241, 1101242-1101243, 1101244-1101245, 1101246-1101247, 1101248-1101249, 1101250-1101251, 1101252-1101253, 1101254-1101255, 1101256-1101257, 1101258-1101259, 1101260-1101261, 1101262-1101263, 1101264-1101265, 1101266-1101267, 1101268-1101269, 1101270-1101271, 1101272-1101273, 1101274-1101275, 1101276-1101277, 1101278-1101279, 1101280-1101281, 1101282-1101283, 1101284-1101285, 1101286-1101287, 1101288-1101289, 1101290-1101291, 1101292-1101293, 1101294-1101295, 1101296-1101297, 1101298-1101299, 1101300-1101301, 1101302-1101303, 1101304-1101305, 1101306-1101307, 1101308-1101309, 1101310-1101311, 1101312-1101313, 1101314-1101315, 1101316-1101317, 1101318-1101319, 1101320-1101321, 1101322-1101323, 1101324-1101325, 1101326-1101327, 1101328-1101329, 1101330-1101331, 1101332-1101333, 1101334-1101335, 1101336-1101337, 1101338-1101339, 1101340-1101341, 1101342-1101343, 1101344-1101345, 1101346-1101347, 1101348-1101349, 1101350-1101351, 1101352-1101353, 1101354-1101355, 1101356-1101357, 1101358-1101359, 1101360-1101361, 1101362-1101363, 1101364-1101365, 1101366-1101367, 1101368-1101369, 1101370-1101371, 1101372-1101373, 1101374-1101375, 1101376-1101377, 1101378-1101379, 1101380-1101381, 1101382-1101383, 1101384-1101385, 1101386-1101387, 1101388-1101389, 1101390-1101391, 1101392-1101393, 1101394-1101395, 1101396-1101397, 1101398-1101399, 1101400-1101401, 1101402-1101403, 1101404-1101405, 1101406-1101407, 1101408-1101409, 1101410-1101411, 1101412-1101413, 1101414-1101415, 1101416-1101417, 1101418-1101419, 1101420-1101421, 1101422-1101423, 1101424-1101425, 1101426-1101427, 1101428-1101429, 1101430-1101431, 1101432-1101433, 1101434-1101435, 1101436-1101437, 1101438-1101439, 1101440-1101441, 1101442-1101443, 1101444-1101445, 1101446-1101447, 1101448-1101449, 1101450-1101451, 1101452-1101453, 1101454-1101455, 1101456-1101457, 1101458-1101459, 1101460-1101461, 1101462-1101463, 1101464-1101465, 1101466-1101467, 1101468-1101469, 1101470-1101471, 1101472-1101473, 1101474-1101475, 1101476-1101477, 1101478-1101479, 1101480-1101481, 1101482-1101483, 1101484-1101485, 1101486-1101487, 1101488-1101489, 1101490-1101491, 1101492-1101493, 1101494-1101495, 1101496-1101497, 1101498-1101499, 1101500-1101501, 1101502-1101503, 1101504-1101505, 1101506-1101507, 1101508-1101509, 1101510-1101511, 1101512-1101513, 1101514-1101515, 1101516-1101517, 1101518-1101519, 1101520-1101521, 1101522-1101523, 1101524-1101525, 1101526-1101527, 1101528-1101529, 1101530-1101531, 1101532-1101533, 1101534-1101535, 1101536-1101537, 1101538-1101539, 1101540-1101541, 1101542-1101543, 1101544-1101545, 1101546-1101547, 1101548-1101549, 1101550-1101551, 1101552-1101553, 1101554-1101555, 1101556-1101557, 1101558-1101559, 1101560-1101561, 1101562-1101563, 1101564-1101565, 1101566-1101567, 1101568-1101569, 1101570-1101571, 1101572-1101573, 1101574-1101575, 1101576-1101577, 1101578-1101579, 1101580-1101581, 1101582-1101583, 1101584-1101585, 1101586-1101587, 1101588-1101589, 1101590-1101591, 1101592-1101593, 1101594-1101595, 1101596-1101597, 1101598-1101599, 1101600-1101601, 1101602-1101603, 1101604-1101605, 1101606-1101607, 1101608-1101609, 1101610-1101611, 1101612-1101613, 1101614-1101615, 1101616-1101617, 1101618-1101619, 1101620-1101621, 1101622-1101623, 1101624-1101625, 1101626-1101627, 1101628-1101629, 1101630-1101631, 1101632-1101633, 1101634-1101635, 1101636-1101637, 1101638-1101639, 1101640-1101641, 1101642-1101643, 1101644-1101645, 1101646-1101647, 1101648-1101649, 1101
```

## ORIGIN

Query Match	58.7%	Score	747.4	DB.1	Length	778			
Best Local Similarity	97.7%	Pred. No.	1.6e-171						
Matches	757	Conservative	0	Mismatches	18	Indels	0	Gaps	0
Qy	159	TTTTATTTTTTTTAAATATCTTTTTTAAATCCCTATTTTCTCTCTCTTTTGCCTCAGTAAATTT	218						
Db	4	TTTTATTTTTTTTAAATATCTTTTTTAAATCCCTATTTCTTCTCTCTTTTGCCTCAGTAAATTT	63						
Qy	219	TGTAATGAACCTTTAAAAGGACTTAAGGATGTAAACATTAATTATTTAAAGTAAGTCATGGT	278						
Db	64	TGTAATGAACCTTTAAAAGGACTTAAGGATGTAAACATTAATTATTTAAAGTAAGTCATGGT	122						
Qy	279	TATAATTAATTTTCTCTGCTCCTCTATGTATTAATTATTTTCAAAATAGCCCTCAGACTGC	338						
Db	124	TATAATTAATTTTCTCTGCTCCTCTATGTATTAATTATTTTCAAAATAGCCCTCAGACTGC	183						
Qy	339	TACAGCATTACTACCGGTACTTCGAAAGTGTCCACCATCCAGAGGGTGCTGCCTCGAC	398						
Db	184	TACAGCATTACTACCGGTACTTCGAAAGTGTCCACCATCCAGAGGGTGCTGCCTCGAC	243						
Qy	399	TGGCAAACTCGATCCAAACAATGACTTGGCGAGTCTTTTTCAGTGTCCAGTCTCGTTGA	458						
Db	244	TGGCAAACTCGATCCAAACAATGACTTGGCGAGTCTTTTTCAGTGTCCAGTCTCGTTGA	303						
Qy	459	CTATGTGTACCGCCATTCCTCAATGTCCAGTGGCCATCTTGTGTGTAGCAACTGTGCG	518						
Db	304	CTATGTGTACCGCCATTCCTCAATGTCCAGTGGCCATCTTGTGTGTAGCAACTGTGCG	363						
Qy	519	CCCAAGCTCAGATGTGTCCAACTTGC CGGGGCCCTTTGGGATCCATTCGCAACTTGGC	578						
Db	364	CCCAAGCTCAGATGTGTCCAACTTGC CGGGGCCCTTTGGGATCCATTCGCAACTTGGC	423						
Qy	579	TATGGAGAAAGTGGCTAATTGAGTACTTTTCCCCTGTAATAATCGCTCTCTCGGATGTA	638						
Db	424	TATGGAGAAAGTGGCTAATTGAGTACTTTTCCCCTGTAATAATCGCTCTCTCGGATGTA	483						
Qy	639	AATRACTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCTTA	698						
Db	484	AATRACTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCTTA	543						
Qy	699	TTCTGTCGTCGCTCGTGGTCTTCTGTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCC	758						

544	TTCTGTGCGTGGCCCTGGTGCTTCTCTGTAAATGGCAAGGCTCTCTGGAATCTCTTAATGCC	603
Db		
759	CCATCTGATGCATCAGCATTAAGTCCCAATTACAACCCCTACAGGAGAGATATAGTTTTCT	818
Qy		
604	CCATCTGATGCATCAGCATTAAGTCCCAATTACAACCCCTACAGGAGAGATATAGTTTTCT	663
Db		
819	TGCTACAGACATTAATCTTCCCTGGTGCTGTGACTGGGTGATGATGACGTCCTGTTTTGG	878
Qy		
564	TGCTACAGACATTAATCTTCCCTGGTGCTGTGACTGGGTGATGATGACGTCCTGTTTTGG	723
Db		
879	CTTTTCATCTCATGTTAGTCTTAGAGAAACAGGAAATATACGTTGGTCACCAGCAG	933
Qy		
724	CTTTTCATCTTCATGTTAGTCTTAGAGAAACAGGAAATATACGTTGGTCACCAGCAG	778
Db		

RESULT 13

[illegible]

ACCESSION

accession  
 version  
 keywords  
 source  
 organism  
 reference  
 authors  
 title  
 journal  
 comment

BX448317.2  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 872)  
 Li, W.B., Gruber, C., Jesses, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 22, 2003 this sequence version replaced gi:31025931.

111

Genoscope - Centre National du Séquençage  
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(3T) primer. Five primer  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 8104.r  
For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?s=CS0AM009AD120P1&c=8104.i>.

**ΕΡΛΑΤΙΠΡΕΣ**

```

1. .872
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CSODM009G23"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone.lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCWSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCWSPORT 6
vector. Library was not normalized."

```

## ORIGIN

	Query Match	57.5%;	Score 733;	DB 5;	Length 872;	
	Best Local Similarity	97.3%;	Pred. No. 5.4e-168;			
	Matches 750;	Conservative 14;	Mismatches 5;	Indels 2;	Gaps 2	
QY	317	CAGAAATGAGCGGTGAGACTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACCAT	376			
DB	1	CAGAAATGAGCGGTGAGACTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACCAT	60			
QY	377	CCGAGAGGGTGCTGCTGCTGATCTGGCACTGCTGATCGAACATGACTTGGCGAGTCT-T	435			
DB	61	CCGAGAGGGTGCTGCTGCTGATCTGGCACTGCTGATCGAACATGACTTGGCGAGTCTAT	120			



```
QY 436 TTGTGAGTGTCCAGTCTGCTTTGACTATGTTTACCGCCCAATTTCTCAATGTCAGAGTGGC 495
Db 121 TTTGAATATCCAGTCTGCTTTGACTATGTTTACCGCCCAATTTCTCAATGTCAGAGTGGC 180
QY 496 CATCTTTGTTGTAGCAACTCTGCGCCCAAGCTCACATGTTGTCAACTTGCCTGGGGCCCT 555
Db 181 CATCTTTGTTGTAGCAACTCTGCGCCCAAGCTCACATGTTGTCAACTTGCCTGGGGCCCT 240
QY 556 TTGGGATCCATTCGCAACTTGGCTATGGAAGAGTGCTAATTCAGTACTTTTCCCTGT 615
Db 241 TTGGGATCCATTCGCAACTTGGCTATGGAAGAGTGCTAATTCAGTACTTTTCCCTGT 300
QY 616 AATATGCGTCTCTGATGCTGAATTAATCTGCCACACACAGAAAAGCAGACCATGAA 675
Db 301 AATATGCGTCTCTGATGCTGAATTAATCTGCCACACACAGAAAAGCAGACCATGAA 360
QY 676 GAGCTCTGTGAGTTTAGGCTTATTTCTGTCCTGCGCTGCTGCTTCTGTAATGCGCA 735
Db 361 GAGCTCTGTGAGTTTAGGCTTATTTCTGTCCTGCGCTGCTGCTTCTGTAATGCGCA 420
QY 736 GGCTCTCTGATGCTGTAATGCGCTTCTGATGATGATGATGATGATGATGATGATGATGAT 795
Db 421 GGCTCTCTGATGCTGTAATGCGCTTCTGATGATGATGATGATGATGATGATGATGATGAT 480
QY 796 CAGGAGAGGATATGTTTTCTTGTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTG 855
Db 481 CAGGAGAGGATATGTTTTCTTGTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 856 GTGATGATGAGTCTGTTTGTGCTTCTCACTTCATGTTAGTCTTAGAGAAACAGGAAAA 915
Db 541 GTGATGATGAGTCTGTTTGTGCTTCTCACTTCATGTTAGTCTTAGAGAAACAGGAAAA 600
QY 916 TAGATGAGTACAGCAGTCTTTCGCAATCGTACAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 975
Db 601 TAGATGAGTACAGCAGTCTTTCGCAATCGTACAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 660
QY 976 GAAATTTTGTCTTACGACTGTAGCTTAATGCTGATGAGGAGGATGCTTGGGAAGCG 1035
Db 661 GAAATTTTGTCTTACGACTGTAGCTTAATGCTGATGAGGAGGATGCTTGGGAAGCG 720
QY 1036 ACTCTCTGATCTTATCATGAAGAAATGCAACAGCATTATGATGAAGCAG 1086
Db 721 MYCCTCGTATTCATGAAGAAATGCAACAGCATTATGATGAAGCAG 770

BX384656 785 bp mRNA linear EST 26-APR-2004
BX384656 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK012Y120 5-PRIME, mRNA sequence.
BX384656
BX384656.2 GI:46570911
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 785)
Li, W.-B.; Gruber, C.; Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30447286.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8104.r
For more information about this cluster, see
```

```
http://www.genoscope.cns.fr/cdna?s=CS0DK012BE10QPl&c=8104.r.
Location/Qualifiers
1..785
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK012Y120"
/cell_type="HELA HELA"
/cell_line="HELA"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 56.4%; Score 719; DB 5; Length 785;
Best Local Similarity 93.8%; Pred. No. 1.4e-164;
Matches 738; Conservative 26; Mismatches 21; Indels 2; Gaps 2;
QY 317 CAGAAATGAGCCGTGAGCTGCTTACAGCATTACTTACCGGTACTCGAAGTGTCCACCAT 376
Db 1 CAGAAATGAGCCGTGAGCTGCTTACAGCATTACTTACCGGTACTCGAAGTGTCCACCAT 60
QY 377 CCCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
Db 61 CCCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 437 TTGAGTGTCCAGTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 496
Db 121 TTGAGTGTCCAGTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 497 ATCTTGTGTTGAGCAACTGTGCGCCCAAGCTCACATGTTGTGCTGCTGCTGCTGCTGCTGCTG 556
Db 181 ATCTTGTGTTGAGCAACTGTGCGCCCAAGCTCACATGTTGTGCTGCTGCTGCTGCTGCTGCTG 240
QY 557 TGGGATCCATTCGCAACTTGGCTATGGAAGAGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 616
Db 241 TGGGATCCATTCGCAACTTGGCTATGGAAGAGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 617 AATATGCGTCTCTGAGTGTGAATTAATCTGCGCACACAGAAAAGCAGACCATGAAG 676
Db 301 AATATGCGTCTCTGAGTGTGAATTAATCTGCGCACACAGAAAAGCAGACCATGAAG 360
QY 677 AGCTCTGTGAGTTTAGGCTTATTTCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736
Db 361 AGCTCTGTGAGTTTAGGCTTATTTCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 737 GCTCTCTGATGCTGTAATGCGCCCATCTGATGATGATGATGATGATGATGATGATGATGATG 796
Db 421 GCTCTCTGATGCTGTAATGCGCCCATCTGATGATGATGATGATGATGATGATGATGATGATG 480
QY 797 AGGGAGAGGATATGTTTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
Db 481 AGGGAGAGGATATGTTTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 857 TGATGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916
Db 541 TGATGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 917 ACATGCTGATGCTGTAATGAGGAATTTGCAACAGCATTATGATGATGATGATGATGATGATG 976
Db 601 ACATGCTGATGCTGTAATGAGGAATTTGCAACAGCATTATGATGATGATGATGATGATGATG 660
QY 1037 CTCCTCGATCTTATCATGAAGAAATTTGCAACAGCATTATGATGATGATGATGATGATGATG 1096
Db 720 CTCCTCGATCTTATCATGAAGAAATTTGCAACAGCATTATGATGATGATGATGATGATGATG 778
QY 1097 TTGACAC 1103
```



617	QY	AAATATGCGCTTCTCGAGTGTGAATAACTCTGCCACACACAGAAAGCAGACCATGAAG	676
301	Db	AAATATGCGCTTCTCGAGTGTGAATAACTCTGCCACACAGAAAGCAGACCATGAAG	360
677	QY	AGCTCTGTGAGTTTATAGCCCTTATCTCTGTCCGTGCCCTGGTGTCTCTCTAAATGGCAAG	736
361	Db	AGCTCTGTGAGTTTATAGCCCTTATCTCTGTCCGTGCCCTGGTGTCTCTCTAAATGGCAAG	420
737	QY	GCTCTCTGGATGCTGTAAATGCCCCATCTGATGCATCAGCATTAAGTCCATTACAACCCCTAC	796
421	Db	GCTCTCTGGATGCTGTAAATGCCCCATCTGATGCATCAGCATTAAGTCCATTACAACCCCTAC	480
797	QY	AGGGAGAGAGATAGTTTTTTCTTGCTACAGACATTAATCTTCTGTGTGCTGTTTGACTGGG	856
481	Db	AGGGAGAGAGATAGTTTTTTCTTGCTACAGACATTAATCTTCTGTGTGCTGTTTGACTGGG	540
857	QY	TGATGATGCAGTCCCTGTTTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAAAT	916
541	Db	TGATGATGC-GTTCTGTTTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAAAT	599
917	QY	ACGATGGTCACAGCAGTTCTTTCGCAATCGTACAGCTGATAGGAACACCCACGACGAGCTG	976
600	Db	ACGATGGTCACAGCAGTTCTTTCGCAATCGTACAGCTGATAGG-AWACGCAAGCMAGCTG	658
977	QY	AAAAATTTTCTCTTACCGACTTGAGCTAAATGGTGCATAGGCGACGATTGACTTTGGGAACCGA	1036
659	Db	AAAAATTTTCTCTTACCGACTTGAGCTAAATGGTGCATAGGCGCGTGDTTTTNTGGGRACGA	718
1037	QY	CTCTCCGATCTATTTCATGAAGGAATTGCAACGCCATTTATGAATA	1081
719	Db	MTCTCTGCATCYAWTCAGGAAGG-AWTTGCAAMARCMATAAGRAAAA	762

Search completed: April 24, 2005, 22:53:18  
Job time : 4607 secs

DB	779	TTGACAC	785		
RESULT 15	BX463976				
LOCUS	BX463976 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone CS0DF021YA01 5-PRIME, mRNA sequence.	852 bp	mRNA	linear	EST 05-MAY-2004
DEFINITION					
ACCESSION	BX463976				
VERSION	1				
KEYWORDS	EST.				
SOURCE	BM463976.2	GI:47053846			
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 852)				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	On May 22, 2003 this sequence version replaced gi:31023592. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8104.r For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CSIAP006ZA010P1&c=8104.r.				

ORIGIN	Query Match	55.6%	Score 708.8;	DB 5;	Length 852;
	Best Local Similarity	96.5%;	Pred. No. 4.4e-162;		
	Matches 738;	Conservative 15;	Mismatches 9;	Indels 3;	Gaps 3;
Qy	317	CAGAAATGAGCGGTGACAGCTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACCAT	376		
Db	1	CAGAAATGAGCGGTGACAGCTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACCAT	60		
Qy	377	CCCAGAGGGTGCCCTGCCTGACTGGCCAACTGCATCCAAATGACTTGGCGAGTCTTT	436		
Db	61	CCCAGAGGGTGCCCTGCCTGACTGGCCAACTGCATCCAAATGACTTGGCGAGTCTTT	120		
Qy	437	TTGAGTGTCCAGTGTGCTTTTGACTATGTGTTACCGCCCATTTCTTCAATGTTCAGAGTGCC	496		
Db	121	TTGAGTGTCCAGTGTGCTTTTGACTATGTGTTACCGCCCATTTCTTCAATGTTCAGAGTGCC	180		
Qy	497	ATCTTTGTTTGTAGCAACTGTGCGCCCAAGCTCACATGTGTGCCAATTCGCGGGGCCCTT	556		
Db	181	ATCTTTGTTTGTAGCAACTGTGCGCCCAAGCTCACATGTGTGCCAATTCGCGGGGCCCTT	240		
Qy	557	TGGATCCATTCGAACTTGCGCTATGGAGAAAGTGGCTAATTCCAGTACTTTTCCCTGTGA	616		
Db	241	TGGATCCATTCGAACTTGCGCTATGGAGAAAGTGGCTAATTCCAGTACTTTTCCCTGTGA	300		

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1274	100.0	1274	4	US-09-591-694-1	Sequence 11, Appli
2	1042.8	81.9	2440	4	US-09-023-655-588	Sequence 586, App
3	928.4	72.9	1884	4	US-09-544-618-11	Sequence 11, Appl
4	161	12.6	257	4	US-09-016-434-397	Sequence 397, App
5	69.6	5.5	1420	3	US-09-362-506-1	Sequence 1, Appli
6	49	3.8	187169	4	US-09-949-016-12776	Sequence 12776, A
7	49	3.8	191569	4	US-09-949-016-15940	Sequence 15940, A
8	48.2	3.8	1425	3	US-09-325-932A-200	Sequence 200, App
9	48	3.8	7218	1	US-08-232-463-14	Sequence 14, Appl
10	47.6	3.7	5852	1	US-07-867-106-2	Sequence 2, Appli
11	47.4	3.7	612	4	US-09-902-540-1357	Sequence 1357, Ap
12	46.8	3.7	39176	4	US-09-949-016-17603	Sequence 17603, A
13	46.6	3.7	1036	3	US-09-325-932A-39	Sequence 39, Appl
14	46.6	3.7	1039	4	US-09-902-540-1280	Sequence 1280, Ap
15	44.8	3.5	832	4	US-09-621-976-2813	Sequence 2813, Ap
16	43.8	3.4	2861	1	US-08-299-953-1	Sequence 1, Appli
17	43.8	3.4	2861	1	US-08-459-415-1	Sequence 1, Appli
18	43.8	3.4	2861	3	US-09-066-687-1	Sequence 1, Appli
19	43.8	3.4	2861	5	PCR-US95-11231-1	Sequence 1, Appli
20	43.8	3.4	3881	1	US-08-299-953-2	Sequence 2, Appli
21	43.8	3.4	3881	1	US-08-459-415-2	Sequence 2, Appli
22	43.8	3.4	3881	3	US-09-066-687-2	Sequence 2, Appli
23	43.8	3.4	3881	5	PCR-US95-11231-2	Sequence 2, Appli
24	43.2	3.4	19124	2	US-08-487-828B-13	Sequence 13, Appl
25	43	3.4	6243	2	US-09-056-075-1	Sequence 1, Appli
26	42.8	3.4	601	4	US-09-949-016-30532	Sequence 30532, A
27	42.8	3.4	601	4	US-09-949-016-30533	Sequence 30533, A

301 CCTATGATATTTATTCAGAAATGAGCGTACAGTCTACAGCAATTAACCTACCGTACC 360  
361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCGCTGACTGGCAGCAACTGCATCCAAACAT 420  
361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCGCTGACTGGCAGCAACTGCATCCAAACAT 420  
421 GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTGTTGACTATGTTTACCGCCATTTCT 480  
421 GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTGTTGACTATGTTTACCGCCATTTCT 480  
481 CAATGTCCAGTGGCCATCTTTTGTAGCAACTGTCGCCAAAGCTCACATGTTGTCCA 540  
481 CAATGTCCAGTGGCCATCTTTTGTAGCAACTGTCGCCAAAGCTCACATGTTGTCCA 540  
541 ACTTGGCGGGGCCCTTTGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATCA 600  
541 ACTTGGCGGGGCCCTTTGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATCA 600  
601 GTACTTTTCCCTGTAAATATGCTTCTGATGTAAGTAATTAATCTGCGCACACAGAA 660  
601 GTACTTTTCCCTGTAAATATGCTTCTGATGTAAGTAATTAATCTGCGCACACAGAA 660  
661 AAAGCAGACATCAAGAGCTCTGTGAGTTTGGCTTATTCCTGTCGGTCCCTGCTGCT 720  
661 AAAGCAGACATCAAGAGCTCTGTGAGTTTGGCTTATTCCTGTCGGTCCCTGCTGCT 720  
721 TCCTGTAATGGCAAGCTCTCTGAGTGTGTAATGCGCCATCTGATGATCAGCATAG 780  
721 TCCTGTAATGGCAAGCTCTCTGAGTGTGTAATGCGCCATCTGATGATCAGCATAG 780  
781 TCATTAACACCTCAGGAGGAGATAGTTTTCTTCTGCTACACACATTAATCTTCT 840  
781 TCATTAACACCTCAGGAGGAGATAGTTTTCTTCTGCTACACACATTAATCTTCT 840  
841 GGTGCTGTGCTGAGTGGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
841 GGTGCTGTGCTGAGTGGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
901 GAGAAACAGGAAATACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
901 GAGAAACAGGAAATACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
961 ACAGCAAGCAAGTGAATTTTCTTACCGACTTGGCTAAATGCTATGCTAGCGCAGCA 1020  
961 ACAGCAAGCAAGTGAATTTTCTTACCGACTTGGCTAAATGCTATGCTAGCGCAGCA 1020  
1021 TTGACTTGGGAAGCGACTCTCGATCTAATCATGAAGGAATTCGCAAGCCATTAATGAAT 1080  
1021 TTGACTTGGGAAGCGACTCTCGATCTAATCATGAAGGAATTCGCAAGCCATTAATGAAT 1080  
1081 AGGCACTGTCTAGTCTTTGACAGCAATGTCAGCTTTTTCAGAAATGCGCAATTTA 1140  
1081 AGGCACTGTCTAGTCTTTGACAGCAATGTCAGCTTTTTCAGAAATGCGCAATTTA 1140  
1141 GGCATCAATGTAATTTTCCATGTTGTAATGCGCAATCAAACTTTCTGCGCCAGTGT 1200  
1141 GGCATCAATGTAATTTTCCATGTTGTAATGCGCAATCAAACTTTCTGCGCCAGTGT 1200  
1201 TTAACACTTCAGTTTTCAGAAATAAGGCAACCATCTGCTGCGCAACCTTAAACCTTTT 1260  
1201 TTAACACTTCAGTTTTCAGAAATAAGGCAACCATCTGCTGCGCAACCTTAAACCTTTT 1260  
1261 CGGTAGTGGNAGC 1274  
1261 CGGTAGTGGNAGC 1274

RESULT 2  
US-09-023-655-588  
; Sequence 588, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.  
APPLICANT: Susan E. Stuart  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,655  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 588:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2440 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: SPINPET01  
CLONE: 29244  
US-09-023-655-588  
Query Match 81.9%; Score 1042.8; DB 4; Length 2440;  
Best Local Similarity 99.8%; Pred. No. 1.1e-296;  
Matches 1044; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 229 TTTAAAGGACTTATGGCATGTAAACATTTTATATAAGTAAGTCATGTTTATTAATT 288  
DB 462 TTTGGAAGGACTTATGGCATGTAAACATTTTATATAAGTAAGTCATGTTTATTAATT 521  
QY 289 TTTCTCTCCCTCCTTATGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 348  
DB 522 TTTCTCTCCCTCCTTATGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 581  
QY 349 CCTACCGGTACCTCGAAGTGTCCACATCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCT 408  
DB 582 CCTACCGGTACCTCGAAGTGTCCACATCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCT 641  
QY 409 GCATCCAAACAATGACTTGGGAGTCTTTTGTAGTGTCCAGTCTGCTTTCGACTATGTTT 468  
DB 642 GCATCCAAACAATGACTTGGGAGTCTTTTGTAGTGTCCAGTCTGCTTTCGACTATGTTT 701  
QY 469 CCGCCCAATTTTCAATGTTCAGAGTGGCCATCTTTTGTAGCAACTGTGCTGCTGCTGCTGCTGCT 528  
DB 702 CCGCCCAATTTTCAATGTTCAGAGTGGCCATCTTTTGTAGCAACTGTGCTGCTGCTGCTGCTGCT 761  
QY 529 ACATGTTGTCACAACTTGGCGGGCCCTTTTGGGATCCATTCGCAACTTGGCTATGGAGAAA 588  
DB 762 ACATGTTGTCACAACTTGGCGGGCCCTTTTGGGATCCATTCGCAACTTGGCTATGGAGAAA 821  
QY 589 GTGGCTAAATTCAGTACTTTTCCCTCTTAAATATGCGTCTTCTGATGTGAAATTAATCTGCT 648

Db 822 GTGGCTAAATTCAGTACTTTTCCCTGTAATATGCGTCTTCGATGTGAATAAATCTCTG 881  
Qy 649 CCACACACAGAAAAGCAGACCAATGAAGAGCTCTGTGAGTTTGGCTTATTCCTGTCCG 708  
Db 882 CCACACACAGAAAAGCAGACCAATGAAGAGCTCTGTGAGTTTGGCTTATTCCTGTCCG 941  
Qy 709 TGCCCTGGTGTCTTCTGTAATGCGAAGGCTCTCTGATGCTGTATGCCCCATCTGATG 768  
Db 942 TGCCCTGGTGTCTTCTGTAATGCGAAGGCTCTCTGATGCTGTATGCCCCATCTGATG 1001  
Qy 769 CATCAGCATAAATGCCATTACACCCCTACAGGAGAGATATAGTTTCTGTGTCAGAC 828  
Db 1002 CATCAGCATAAATGCCATTACACCCCTACAGGAGAGATATAGTTTCTGTGTCAGAC 1061  
Qy 829 ATTAATCTTCTGGTGTCTGTGATGCGGTGATGATGAGTCTCTGTGTTTGGCTTTCCTTC 888  
Db 1062 ATTAATCTTCTGGTGTCTGTGATGCGGTGATGATGAGTCTCTGTGTTTGGCTTTCCTTC 1121  
Qy 889 ATGTAGTCTTAGAGAAAACAGGAAAATACGATGGTTCACAGAGTCTTCGCAATCGTA 948  
Db 1122 ATGTAGTCTTAGAGAAAACAGGAAAATACGATGGTTCACAGAGTCTTCGCAATCGTA 1181  
Qy 949 CAGCTCATAGGAACACGCAAGCAAGCTGAAAATTTTGGCTTACCGACTTGAGCTTAATGTT 1008  
Db 1182 CAGCTCATAGGAACACGCAAGCAAGCTGAAAATTTTGGCTTACCGACTTGAGCTTAATGTT 1241  
Qy 1009 CATAGCGAGATGACTTTGGGAAGCGACTCTCTCGATCTATTCATGAAGGAATTCGCAACA 1068  
Db 1242 CATAGCGAGATGACTTTGGGAAGCGACTCTCTCGATCTATTCATGAAGGAATTCGCAACA 1301  
Qy 1069 GGCATTATGAATAGCAGTCTGTAGTCTTTGACACAGAGATTCGACAGCTTTTGGCAGAA 1128  
Db 1302 GGCATTATGAATAGCAGTCTGTAGTCTTTGACACAGAGATTCGACAGCTTTTGGCAGAA 1361  
Qy 1129 AATGGCAATTTAGGCATCAATGTAATTTTCCATGTTTGAATGGCAATCAACATTT 1188  
Db 1362 AATGGCAATTTAGGCATCAATGTAATTTTCCATGTTTGAATGGCAATCAACATTT 1421  
Qy 1189 TCTGGCCAGTGTTTAAAACCTTCAGTTTTCAGAAAATAAGGCACCCATCTGTCTGCCAAC 1248  
Db 1422 TCTGGCCAGTGTTTAAAACCTTCAGTTTTCAGAAAATAAGGCACCCATCTGTCTGCCAAC 1274  
Qy 1249 CTAAAACCTCTTCGGTAGGTGGAAGC 1274  
Db 1482 CTAAAACCTCTTCGGTAGGTGGAAGC 1507

## RESULT 3

US-09-544-618-11  
; Sequence 11, Application US/09544618  
; Patent No. 6503502  
; GENERAL INFORMATION:  
; APPLICANT: Teleman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Cohen, Daniel  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, PROTEINS, DRUGS AND DIAGNOSTIC  
; TITLE OF INVENTION: AGENTS OF USE IN TREATING CANCER  
; FILE REFERENCE: 065691-0139  
; CURRENT APPLICATION NUMBER: US/09/544,618  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 1884  
; TYPE: DNA  
; ORGANISM: TSAP 3  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(846)  
US-09-544-618-11

Query Match

72.9%; Score 928.4; DB 4; Length 1884;

Best Local Similarity 99.7%; Pred. No. 4.9e-263;  
Matches 951; Conservative 0; Mismatches 1; Indels 2; Gaps 2;  
Qy 322 ATGAGCCGTGAGACTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACCATCCCG 381  
Db 1 ATGAGCCGTGAGACTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACCATCCCG 60  
Qy 382 AGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441  
Db 61 AGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
Qy 442 TGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501  
Db 121 TGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
Qy 502 GTTGTGTAGCAATCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561  
Db 181 GTTGTGTAGCAATCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
Qy 562 TCCATTCGCAACTTTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTAATAT 621  
Db 241 TCCATTCGCAACTTTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTAATAT 300  
Qy 622 GCGTCTTCTGGATGTGAATTAATCTCTGCCACACACAGAAAAGCAGACCATGAGAGCTC 681  
Db 301 GCGTCTTCTGGATGTGAATTAATCTCTGCCACACACAGAAAAGCAGACCATGAGAGCTC 360  
Qy 682 TGTGAGTTTGGCTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 741  
Db 361 TGTGAGTTTGGCTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
Qy 742 CTGGATGTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801  
Db 421 CTGGATGTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Qy 802 GAGGATATAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 861  
Db 481 GAGGATATAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Qy 862 ATGCACTCTGTTTGGCTTTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 921  
Db 541 ATGCACTCTGTTTGGCTTTCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Qy 922 GGTCAACAGCAGTCTTCTGCAATCTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTG 981  
Db 601 GGTCAACAGCAGTCTTCTGCAATCTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTG 660  
Qy 982 TTTGCTTACCGACTTGGCTAAATGGTCTATAGGCAAGTGTGCTGCTGCTGCTGCTGCTG 1041  
Db 661 TTTGCTTACCGACTTGGCTAAATGGTCTATAGGCAAGTGTGCTGCTGCTGCTGCTGCTG 720  
Qy 1042 CGATCTATTTCATGAAGGAATTCGCAAGCATTATGTAATGAGCACTGTCTAGTCTTTG 1101  
Db 721 CGATCTATTTCATGAAGGAATTCGCAAGCATTATGTAATGAGCACTGTCTAGTCTTTG 780  
Qy 1102 ACCAGATTGCAAGCTTTTGGCAGAAATGCAAGTGTGTAAGTGTGTAAGTGTGTAAGT 1160  
Db 781 -CCAGATTGCAAGCTTTTGGCAGAAATGCAAGTGTGTAAGTGTGTAAGTGTGTAAGT 839  
Qy 1161 CATGTTTGAATGCAATCAACATTTCTGCGCAGTGTGTTAAACTTCAGTTTTCAGTTC 1220  
Db 840 CATGTTTGAATGCAATCAACATTTCTGCGCAGTGTGTTAAACTTCAGTTTTCAGTTC 899  
Qy 1221 AAAATAAGGCACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1274  
Db 900 AAAATAAGGCACCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 953

## RESULT 4

US-09-016-434-397  
; Sequence 397, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:

```

; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 397:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT03
; CLONE: 2113436
; US-09-016-434-397

Query Match      12.6%; Score 161; DB 4; Length 257;
Best Local Similarity 76.7%; Pred. No. 2.7e-37;
Matches 197; Conservative 0; Mismatches 60; Indels 0; Gaps. 0;

QY      629  CTGGATGTGAATAAATCTCTGCCACACACAGAAAAGCAGACCATGAGAGCTCTGTGAGT 688
DB      1   CGGGCTGTTCCTGACCTGCACCATACGGAGAAACCCAGAACATGAAGACATATGTGAAT 60

QY      689  TTAGGCCCTTATCTGTCTGCGTGGCTTCTCTGTAATGCAAGGCTCTCTGGATG 748
DB      61  ACCGTCCCTACTCTGTGCCAATGCTCTGCTGCTTCTCTGCAAGTGGCAGGGTCTCTGGAAG 120

QY      749  CTGTAATGCCCATCTGATGCATCAGCATAAAGTCCATTACAAACCTACAGGGAGAGATA 808
DB      121  CTGTGATGTCCCATCTCATGACGCCACAGAGCATTACACCTTCAGGGAGAGACA 180

QY      809  TAGTTTTTTTGTGTACAGACATTAATCTTCTGTGTGCTGTGTGAGTGGGTGATGATGCAGT 868
DB      181  TCGTCTTTCTAGCTACAGACATTAATCTTGCAGGGGCTGTGCGAGTGGGTGATGATGCAGT 240

QY      869  CCTGTTTGGCTTTCAC 885
DB      241  CATGTTTGGCCATCAC 257

RESULT 5
US-09-362-506-1
; Sequence 1, Application US/09362506
; Patent No. 611167
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12776
; LENGTH: 187169
; TYPE: DNA

; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize SINA Orthologue-1 and Uses Thereof
; FILE REFERENCE: 0936
; CURRENT APPLICATION NUMBER: US/09/362,506
; CURRENT FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: 60/100,258
; EARLIER FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (202)...(1140)
; US-09-362-506-1

Query Match      5.5%; Score 69.6; DB 3; Length 1420;
Best Local Similarity 51.4%; Pred. No. 6.7e-10;
Matches 187; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

QY      420  TGACTTGGCGAGTCTTTTGGAGTGCAGTCTGCTTTGACTATGTGTACCGCCCATTTCT 479
DB      366  TGGCTTGAATGATTTGCTCGAATGCCAGTGTGTACCAACTCGATGCCGCCACCTATACT 425

QY      480  TCAATGTCAGAGTGGCCATCTTTGTTGTAGCAACTGTGCGCCCAAGCT---CACATGTTG 536
DB      426  CCAGTGGCCCAATGGCCACACGATCTGCTAGTTGCAAGCACAGGGGTAGAGAACCATTTG 485

QY      537  TCCAACTTGGCGGGGCCCTTTGGGATCCATTGCGCAACTTGGCTATGGAGAAAGTGGCTAA 596
DB      486  CCCAACTGTGCGCCAGGAACATCGGGAAACATCAGGTGCTCTCGAGAAAGGTGGCAGA 545

QY      597  TTCAGTACTTTTCCCTCTGTAATATGCGTCTTCTGGATGTGAAATAAATCTCTGCCACACAC 656
DB      546  GCAACTCCAGCTTCCGTGCAAGTACAGACAGCAGCGGATGCCACCGAGATCCACCCATACAA 605

QY      657  AGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTGAAGCTTATTCCTGTGCGTGCCTGG 716
DB      606  GAGCAAACTGAAGCAGCAGGAGCTCTGCAGGTTTCAGGCGCTACAGCTGCCCGTACGCAGG 665

QY      717  TGTCTCTGTAATGCGAAGGCTCTCTGGATGCTGTAATGCCCATCTGTATGATGCATCAGCA 776
DB      666  TTCCGAGTGCCTGATGCGCAGGCGACGTCCTGCTGTTCCCACTCATCAACGACCA 725

QY      777  TAAG 780
DB      726  TAAG 729

RESULT 6
US-09-949-016-12776/c
; Sequence 12776, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12776
; LENGTH: 187169
; TYPE: DNA
```

```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(187169)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12776

Query Match      3.8%; Score 49; DB 4; Length 187169;
Best Local Similarity 48.4%; Pred. No. 0.015;
Matches 136; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 10 TTGTTATGGTCATTTCTATTTTGGCAATTTTATTTAGCAATTTATTTCTATGCTAGTCTATCCAAAGACG 69
DB 883 TTATATATTTTATATATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTAT 824

QY 70 ATTAAGGAGTTCCACATGTTTCCGAAACATTTTGAAAGAGAGCTTATCCAGTGTACA 129
DB 823 ATTTTATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTAT 764

QY 130 GATCCTAAATAAGTGACATTCAGTGTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 189
DB 763 TATTATATATTTTATATTTTATATATATTTTATATATTTTATATATTTTATATATTTTATATATTTAT 704

QY 190 TTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 249
DB 703 TATTATATATTTTATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTAT 644

QY 250 TAAACATTTTATATAAGTAAGTCATGGTTATAAATTTATTTT 290
DB 643 TATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTAT 603

RESULT 7
US-09-949-016-15940/c
; Sequence 15940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15940
; LENGTH: 191569
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(191569)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15940

Query Match      3.8%; Score 49; DB 4; Length 191569;
Best Local Similarity 48.4%; Pred. No. 0.015;
Matches 136; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 10 TTGTTATGGTCATTTCTATTTTGGCAATTTTATTTAGCAATTTTATTTCTATGCTAGTCTATCCAAAGACG 69
DB 883 TTATATATTTTATATATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTAT 824

QY 70 ATTAAGGAGTTCCACATGTTTCCGAAACATTTTGAAAGAGAGCTTATCCAGTGTACA 129
DB 823 ATTTTATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTAT 764
```

```
QY 130 GATCCTAAATAAGTGACATTCAGTGTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 189
DB 763 TATTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTATA 704

QY 190 TTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 249
DB 703 TATTATATATATTTTATGTAATTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTATA 644

QY 250 TAAACATTTTATATAAGTAAGTCATGGTTATAAATTTATTTT 290
DB 643 TATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTTATATATTTAT 603

RESULT 8
US-09-325-932A-200
; Sequence 200, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; death and their use in the modification of forestry plant develop
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 200
; LENGTH: 1425
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-200

Query Match      3.8%; Score 48.2; DB 3; Length 1425;
Best Local Similarity 43.4%; Pred. No. 0.0014;
Matches 273; Conservative 0; Mismatches 353; Indels 3; Gaps 1;

QY 528 CACATGTTTGTCCAACCTGCGGGGGCCCTTTGGGATCCATTCCGCAACTTGGCTATGAGAA 587
DB 341 CAACTGGTGTCCGCTCTGCAGATGCAACCTTGGAATTTTAGTGCTTAGCTCTCTGAAAC 400

QY 588 AGTGGCTAATTCAGTACTTTTCCCTGTAAATATGGTCTTCTGGATGTGAATAACTCT 647
DB 401 GGAGACATCATCTCAAGAACTTACTTGCATGTATCAAGCTATGTTGTGAGGATATGA 460

QY 648 GCCACACACAGAAAAAGCAGACCATGAGAGCTCTGTGAGTTTAGGCTTTAGGCTTATTCCTGTCC 707
DB 461 TCTTACTACAGTGAATTAAGACATGAAGCTCCTGCAATTTTAGGCCATACAACTGTCC 520

QY 708 GTGCCCTGTGTCTCTGTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGAT 767
DB 521 CTATGCTGGCTCCGAATGCAAGCTAGTTGGAGATATTCCTTTTGGTGGCTCATTTAAG 580

QY 768 GCATCAGCATAGTCCATTAACCCCTACAGGAGGAGGATATAGTTTTTCTGCTACAGA 827
DB 581 AGATGATCACAAAGTTTATATGCAATATAGTTTGCACCTTTTGATCATCGATATGTAAGTC 640

QY 828 CATTAATCTTCTGTGTG---CTGTTGACTGGGTGATGATGTCAGTCTCTGTTGGCTTTCA 884
DB 641 AAATCCACTCGAGGTTGAGATGCTATTTGGATGCCAAGCTGAATCAATGTTTGGGGA 700

QY 885 CTTTCATGTTAGTCTTTAGAGAAAACAGAAAAATACGATGGTCCACGACAGTCTTCTTCCAAAT 944
DB 701 ATTCCTTTGTCTACATTTTGAAGCGTTTCTATTAGACATGGCCCTGTATATATATAGCTTT 760

QY 945 COTCAGCTGATAGGAACACGCAAGCAGCTGAAATTTTGGCTTACCGACTTGAGCTAAA 1004
DB 761 TCTGATTTTTCATGGGAGATGATAATGAAGCTAAAAAATTTAGCTATTTGCTCGAGACTCG 820

QY 1005 TGGTCATAGGCGACCATTTGACCTTGGGAGCGACTCTCGATCTATTCATGAAGGATTCG 1064
DB 821 AGGCAATGTCGAAACACTGATTTGGCATGGGGTCTCTCGAAGCATCAGAGATTTGTCACAG 880
```







Matches	272;	Conservative	0;	Mismatches	354;	Indels	3;	Gaps	1;
528	Qy	CACATGTTGTCCAACTTGC	CGGGGCCCTTTGGGATCCATT	CGCAATTCGGCTATGAGAA	587				
345	Ddb	CAACTGGGTCCGTCTG	CAGATGCAACCTTGGAAATTT	AGTGTCTTAGCTCCTGAAAC	404				
588	Qy	AGTGGCTAATTCAGTACT	TTTCCCTGTGTAATATGCGTCTT	CTGGATGTGAATACTCT	647				
405	Ddb	GGAGACATCATCTCAAGA	ACTTACTTTGCATGTATCAA	AGCTATGGTGTGAGGATATGA	464				
648	Qy	GCCACACACAGAAAAAG	CAGACCATGAAGAGCTCTGT	GAGTCTGTAGGCTTATTCCTGTGTC	707				
465	Ddb	TCCTTACTACAGTGAAT	TAAAGACATGAAGCTCAGT	GCCAAATTTAGGCCATACAACTGTCC	524				
708	Qy	GTGCCCTGGTGTCTCTG	TAAATGGCAAGCTCTCTG	GGATCTGTGTAATGCCCATCTGAT	767				
525	Ddb	CTATGCTGGCTCGAATG	CAAGCTAGTTGGAGATATTC	CTTTTGGTGGCTCATTTAAG	584				
768	Qy	GCATCAGCATAGTCCAT	TACAACTACAGGAGAGGATAT	AGTTTTTCTTGCTACAGA	827				
585	Ddb	AGATGATCAAAAGTTT	TATGCAATAATAGTTGCAC	CTTTGATCATCGATGTAAGTC	644				
828	Qy	CATTAATCTTCTCGTG	G---CTGTTCAGCTGGGTGATG	TCAGTCCCTGTTTTGGCTTTCA	884				
645	Ddb	AAATCCACTCGAGTTT	GAGAAATGCTATTTTGGATG	CCAACTGTAAATCAATTTGTTTGGGCA	704				
885	Qy	CTTCATGTTAGTCTTAG	AGAAACAGGAAAAATACGAT	GGTCACCGACAGTTCCTTCGCAAT	944				
705	Ddb	ATTCTTTTGTCTACAT	TTTGAAGCGTTCTATTAGAC	ATGGCCCTGTATATATAGCTTT	764				
945	Qy	CGTACAGCTGATAGGA	ACACGCAAGCTGAAAATTTT	GTCTACCGACTTCAGCTTAAA	1004				
765	Ddb	TCGTATTTTCATGGG	AGATGATAATGAAGCTAA	AAACTTTTAGCTATTGCTCGAGACTGG	824				
1005	Qy	TGCTCATAGCGCAGAT	TGACTTTGGGAGCGACTCCT	CGATCTATTTCATGAAGGAATTC	1064				
825	Ddb	AGGCAATGTCGGAAC	ATGATTTGGCATGGGTTCCT	CGAAGCATCAGAGATTGTCAAG	884				
1065	Qy	AACAGCCATTATGAAT	AGCACTGTCTAGTCTTTG	ACACCGACTTGCACAGCTTTTTTC	1124				
885	Ddb	GAAAGTTTCATGAC	GTAGTAGCGGACTTAAT	TATACAAAGAGATGTGCACTCTTTTCTC	944				
1125	Qy	AGAAATGGCAATTTT	TAGGCATCAATGTAA	1153					
945	Ddb	AGGTGGTGCATAAAT	GAAATGAAATGAACTTA	973					

```

RESULT 14
US-09-902-540-1280/c
; Sequence 1280, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15949)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1039)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

```

		Query Match	3.7%; Score 46.6; DB 4;	Length 1039;
		Best Local Similarity	49.4%;	Pred. No. 0.0034;
Matches 154;	Conservative	0;	Mismatches 150;	Indels 8; Gaps 1;
Qy	5	T T T A G T T G T T A T G C T C C A T T T T C T A T T T T T A G C A T T T A T T A T T C T A T G T A G T C T A T C C A A	64	
Dd	799	T T T A T A T T T A T A T A T T T T T T T T T T T T T T G T T T T T T A G A T T T T T T G T T A T A T T G G G G T C T G C A T T	740	
.Qy	65	A G A C G A T T A A G G A G A T T C C A C A T G T T T T C C G G A A C A T T T T G A A A A G A G A G C T T A T C C A G T	124	
Dd	739	A G A G C C T T T G T A A T T T C T T T T T T T T T T T T T T T T T T A T G T T T T T A G G G G A G A T T T T T T T T T T	680	
Qy	125	G T A C A G A T C C T A A T A A G T G C A C A T T C A G T G T A A T T T T T T T T T T T A A T A T C T T T T T T T A A	184	
Dd	679	T T A -----A G A A A A A A C A C T T T T T T T G T T T T T T A T A G T T T T A T A T T T T A T T T T A T A T	628	
Qy	185	T C C T A T T T T T C T C C T C T T T T G C T C A G T A A A T T T T G T A T A G A A C T T T A A A A G A C T T A T G	244	
Dd	627	T T T A T G C G G N G T A A A T A T A T G T A T T A C T T T T T T T T T T T T T A	568	
Qy	245	G C A T G T A A A C A T A T T T A T A A A G T A A G T C A T G G T T A T A A T T A T T T T T C T C T G C C T C C T T	304	
Dd	567	T C T T T T T T A T T T T T T T T T T T A T T T G T T T T T T T T T T T T T T A T T A T T T T T A T T T T T T T T T T T	508	
Qy	305	A T G P A T T A T T T	316	
Dd	507	G T T T T C A T T T T	496	

```

RESULT 15
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
; US-09-621-976-2813

```

Query Match	3.5%;	Score 44.8;	DB 4;	Length 832;
Best Local Similarity	9.0%;	Pred. No. 0.01;	Mismatches 146;	Indels 0;
Matches	31;	Conservative 169;	Gaps 0;	
QY	38	ATTATTATTTCATGCTAGTCTATCCAAAGACGATTAAAGGAGGTTCCACATGTTTCCCGGA	97	
DB	3	WYWKVYITWYAKCWTKKWWSYVMYWKWYVMKTYWRWRKKKAWWKYKWTWMTWYWR	62	
QY	98	ACATTTTGAAGAAGAGAGCTTATCCAGTGTACAGATCCCTAAATAAAGTGACATTCAGTGTGA	157	
DB	63	AMWGYYKKAMCRKTKTKKKKKGGYMMWWTGWRRRSYMAWWTWTGTGYAYRSMYWRYYR	122	
QY	158	ATTTTTATTTTTTAAATATCTTTTTTAAATCCTATTTTCTCTCTCTTTTCTCAGTAAAT	217	
DB	123	CWKKKAYYKTTTCYSSKGWTVWKRKKAWTITWWWKTTYWAAATRYMMWCMWTKRWBSAW	182	
QY	218	TGTATGAACACTTAAAGGACCTTATGGCATGTAAACATTTATTTATAAAGTCAATCGG	277	
DB	183	YCWGKAKWSTWRKRSYASARSNRCCYSCSWGAMSKYWRWWRWGWATGAGMKA	242	
QY	278	TTATAATTATTTTTCTCCGTCCTCTTATGTATTTATTTTCAGAAATGAGCGGTCAACTG	337	

Db 243 WRASCMRRKYAGKSTYSKSMWMCWTRSWKICYTKARWTGYCYRKGGWGRGRWYAS 302  
Qy 338 CTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCAGAG 383  
Db 303 KKYMKRMMWMCWARMYRYSTGTGRASMMWRWYTYTMMKWKYAWAR 348

Search completed: April 24, 2005, 22:57:24  
Job time : 248 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 21:36:41 ; Search time 821 Seconds  
(without alignments)  
9420.852 Million cell updates/sec

Title: US-10-679-246-1

Perfect score: 1274

Sequence: 1 ttcttcttggtgtttatggt.....ctcttcttggttagtggaagc 1274

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5633728 seqs, 3035525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1274	100.0	1274	18	US-10-679-246-1
2	1274	100.0	2924	18	US-10-357-930-22959
3	1274	100.0	2924	18	US-10-357-930-25457
4	1274	100.0	2924	18	US-10-357-930-28822
5	1274	100.0	6107	10	US-09-764-891-8070
6	1272.4	99.9	2829	17	US-10-108-260A-269
7	1042.8	81.9	2440	17	US-10-641-643-558
8	429	33.7	975	17	US-10-172-118-1117
9	429	33.7	975	17	US-10-342-887-1117
10	429	33.7	2502	15	US-10-177-293-426
11	427.4	33.5	2240	15	US-10-177-293-424

12	427.4	33.5	2240	18	US-10-717-597-137	Sequence 137, App
13	427.4	33.5	2240	18	US-10-755-889-123	Sequence 123, App
14	379	29.7	1535	9	US-09-925-297-84	Sequence 84, Appl
15	293.8	23.1	360	9	US-09-960-352-10011	Sequence 10011, A
16	191.6	15.0	405	18	US-10-357-930-38440	Sequence 38440, A
17	190.6	15.0	440	9	US-09-783-590-4667	Sequence 4667, Ap
18	188.2	14.8	426	18	US-10-357-930-18669	Sequence 18669, A
19	183.4	14.4	473	18	US-10-357-930-48480	Sequence 48480, A
20	183	14.4	277	18	US-10-357-930-8510	Sequence 8510, Ap
21	161	12.6	257	17	US-10-305-720-397	Sequence 397, App
22	160.2	12.6	411	18	US-10-357-930-13819	Sequence 13819, A
23	156.8	12.3	441	18	US-10-357-930-34958	Sequence 34958, A
24	139	10.9	417	18	US-10-357-930-43808	Sequence 43808, A
25	112	8.8	438	18	US-10-357-930-4650	Sequence 4650, Ap
26	110.2	8.6	2142	17	US-10-108-260A-1336	Sequence 1336, Ap
27	103.8	8.1	188	18	US-10-425-115-63145	Sequence 63145, A
28	90.8	7.1	921	18	US-10-425-115-147280	Sequence 147280,
29	83.2	6.5	1835	18	US-10-437-963-55834	Sequence 55834, A
30	81.4	6.4	226	9	US-09-728-443-159	Sequence 159, App
31	81.2	6.4	1438	18	US-10-437-963-83416	Sequence 83416, A
32	81.2	6.4	1448	18	US-10-437-963-84963	Sequence 84963, A
33	80.6	6.3	1383	18	US-10-324-120-4	Sequence 4, Appl
34	75.6	5.9	1526	17	US-10-424-599-21498	Sequence 21498, A
35	74.8	5.9	1342	18	US-10-425-115-183094	Sequence 183094,
36	73.6	5.8	2078	17	US-10-424-599-21496	Sequence 21496, A
37	73.4	5.8	984	9	US-09-938-842A-463	Sequence 463, App
38	73.4	5.8	984	11	US-09-938-842A-463	Sequence 463, App
39	70.6	5.5	165	9	US-09-960-352-1045	Sequence 1045, Ap
40	69.6	5.5	1408	17	US-10-425-114-2756	Sequence 2756, Ap
41	69.6	5.5	1522	18	US-10-425-115-182305	Sequence 182305,
42	68.4	5.4	1398	17	US-10-424-599-98299	Sequence 98299, A
43	67.8	5.3	447	11	US-09-732-627A-4794	Sequence 4794, Ap
44	67	5.3	1378	17	US-10-424-599-15841	Sequence 15841, A
45	66.8	5.2	1683	18	US-10-425-115-27188	Sequence 27188, A

#### ALIGNMENTS

RESULT 1  
US-10-679-246-1  
Sequence 1, Application US/10679246  
Publication No. US20040163138A1  
GENERAL INFORMATION:  
APPLICANT: Reed, John C.  
APPLICANT: Matsuzawa, Shu-ichi  
TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved  
in Protein Degradation, Products and Methods Related Thereto  
FILE REFERENCE: 66821-235  
CURRENT APPLICATION NUMBER: US/10/679,246  
CURRENT FILING DATE: 2003-10-02  
PRIOR APPLICATION NUMBER: US 09/591,694  
PRIOR FILING DATE: 2000-06-09  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 1274  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (274)...(1167)  
US-10-679-246-1

Query Match 100.0%; Score 1274; DB 18; Length 1274;  
Best Local Similarity 100.0%; Pred. No. 2e-301; Mismatches 0; Indels 0; Gaps 0;  
Matches 1274; Conservative 0;  
QY 1 TTCTTCTAGTGTTCATGTCATTTTCTATTTTAGCATTTTATTTCTATGAGTCTAT 60  
Db 1 TTCTTCTAGTGTTCATGTCATTTTCTATTTTAGCATTTTATTTCTATGAGTCTAT 60  
QY 61 CCAAGACGATTAGGGAGTTCACATGTTTTCGGAACATTTTGAAGAGAGCTTATC 120

Db 61 CCAAGAGCATTAAAGGAGTCCACATGTTTCCGGAACATTTGAAAAGAGAGCTTATC 120  
Qy 121 CAGGTACAGATCCTAATAAAGTGCACATTCAGTGTAAATTTTATTTTATATATCTTTT 180  
Db 121 CAGGTACAGATCCTAATAAAGTGCACATTCAGTGTAAATTTTATTTTATATATCTTTT 180  
Qy 181 TTAATCCTATTTTCT 240  
Db 181 TTAATCCTATTTTCT 240  
Qy 241 TATGCGATCTAAACATTTATTAAGTAAGTCAATGTTTATTAATTTTCTCTCTCTCT 300  
Db 241 TATGCGATCTAAACATTTATTAAGTAAGTCAATGTTTATTAATTTTCTCTCTCTCT 300  
Qy 301 CCTTATGTATTTATTTTCAAGATAGCGCTCAGATCTGCTACAGCAATTTACCTACCGGTACC 360  
Db 301 CCTTATGTATTTATTTTCAAGATAGCGCTCAGATCTGCTACAGCAATTTACCTACCGGTACC 360  
Qy 361 TCGAAGTGTCCACCATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 361 TCGAAGTGTCCACCATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Qy 421 GACTGCGAGTCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCCAATTCCT 480  
Db 421 GACTGCGAGTCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCCAATTCCT 480  
Qy 481 CAATGTACAGTGGCCATCTTTTGTAGCACTGTGCGCCCAAGCTCACATGTTGTCCA 540  
Db 481 CAATGTACAGTGGCCATCTTTTGTAGCACTGTGCGCCCAAGCTCACATGTTGTCCA 540  
Qy 541 ACTTGGCGGGCCCTTTGGGATCCATTCGCACTTGGCTATGAGAAAGTGGCTAATTTCA 600  
Db 541 ACTTGGCGGGCCCTTTGGGATCCATTCGCACTTGGCTATGAGAAAGTGGCTAATTTCA 600  
Qy 601 GTACTTTTCCCTGTAAATATGCTCTTCTGGAATGTGAATAAATCTCTGCCACACAGAA 660  
Db 601 GTACTTTTCCCTGTAAATATGCTCTTCTGGAATGTGAATAAATCTCTGCCACACAGAA 660  
Qy 661 AAGCAGACCATGAGAGCTCTGTAGTGTAGGCTTATTCCTGCTGCTGCTGCTGCTGCT 720  
Db 661 AAGCAGACCATGAGAGCTCTGTAGTGTAGGCTTATTCCTGCTGCTGCTGCTGCTGCT 720  
Qy 721 TCCGTAAATGGCAAGCTCTCTGGATGCTGTAAATGCCCATCTGATGCATCAGCATAAG 780  
Db 721 TCCGTAAATGGCAAGCTCTCTGGATGCTGTAAATGCCCATCTGATGCATCAGCATAAG 780  
Qy 781 TCCATTAACAACCTTACAGGAGAGGATATAGTATTTTCTGCTACAGACATTAATCTTCT 840  
Db 781 TCCATTAACAACCTTACAGGAGAGGATATAGTATTTTCTGCTACAGACATTAATCTTCT 840  
Qy 841 GGTGCTGTGACTGGGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Db 841 GGTGCTGTGACTGGGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Qy 901 GAGAAACAGGAAATAACGATGCTTCCAGCAGTCTTCCCAATCGTACAGCTGATAGGA 960  
Db 901 GAGAAACAGGAAATAACGATGCTTCCAGCAGTCTTCCCAATCGTACAGCTGATAGGA 960  
Qy 961 ACAGCAGAGCAAGCTGAAATTTTCTTACCGACTTGAGCTTAATGCTCATAGCGCAGCA 1020  
Db 961 ACAGCAGAGCAAGCTGAAATTTTCTTACCGACTTGAGCTTAATGCTCATAGCGCAGCA 1020  
Qy 1021 TTGACTTGGGAAGCAGCTCTCTCGATCTATTTCATGAAGGAATTCAGCAGCCATTTATGAAT 1080  
Db 1021 TTGACTTGGGAAGCAGCTCTCTCGATCTATTTCATGAAGGAATTCAGCAGCCATTTATGAAT 1080  
Qy 1081 AGCGACTGTCTAGTCTTTTGAACAGCAGCTTTCGAGCTTTTTCGAGAAATGGCAATTTA 1140  
Db 1081 AGCGACTGTCTAGTCTTTTGAACAGCAGCTTTCGAGCTTTTTCGAGAAATGGCAATTTA 1140  
Qy 1141 GGCATCAATGTAACTATTTTCCATGTTGAAATGCAATCAAACTTTCTGGCCAGTGT 1200

Db 1141 GGCATCAATGTAACTATTTTCCATGTTGAAATGGCAATCAAACTTTCTGGCCAGTGT 1200  
Qy 1201 TTAATACTTCAGTTTTCACAGAAATAAGCACCCATCTGTCTGCCAACCTTAAACTCTTT 1260  
Db 1201 TTAATACTTCAGTTTTCACAGAAATAAGCACCCATCTGTCTGCCAACCTTAAACTCTTT 1260  
Qy 1261 CGGTAGGTGGAAGC 1274  
Db 1261 CGGTAGGTGGAAGC 1274  
  
RESULT 2  
US-10-357-930-22959  
; Sequence 22959, Application US/10357930  
; Publication No. US20040259086A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Engede, Wilson  
; APPLICANT: Monahan, John  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
; FILE REFERENCE: MRI-007BCN  
; CURRENT APPLICATION NUMBER: US/10/357,930  
; CURRENT FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: 09/785,276  
; PRIOR FILING DATE: 2003-02-16  
; PRIOR APPLICATION NUMBER: 60/183,319  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: 60/189,862  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/207,454  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/211,314  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/219,007  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/255,281  
; PRIOR FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 62232  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22959  
; LENGTH: 2924  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: 1, 57, 2923, 2924  
; OTHER INFORMATION: n = A,T,C or G  
US-10-357-930-22959  
  
Query Match 100.0%; Score 1274; DB 18; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 3.1e-301; Indels 0; Gaps 0;  
Matches 1274; Conservative 0; Mismatches 0;  
  
Qy 1 TTTCTTTAGTGTGTTATGTTCCATTTTCTATTTTATTTAGCAATTTATTTCTATGCTAT 60  
Db 480 TTTCTTTAGTGTGTTATGTTCCATTTTCTATTTAGCAATTTATTTCTATGCTAT 539  
  
Qy 61 CCAAGACGATTAAGGAGTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120  
Db 540 CCAAGACGATTAAGGAGTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 599  
  
Qy 121 CAGGTACAGATCCTAATAAAGTGCACATTCAGTGTAAATTTTATTTTATATCTTTT 180  
Db 600 CAGGTACAGATCCTAATAAAGTGCACATTCAGTGTAAATTTTATTTTATATCTTTT 659  
  
Qy 181 TTAATCCTATTTTCT 240  
Db 660 TTAATCCTATTTTCT 719  
  
Qy 241 TATGCGATGTAAACATTTATTAAGTAAGTCAATGTTTATTAATTTTCTCTCTCTCT 300

Db 720 TATGGCATGTAACATTAATTATAAAGTAAGTCATGGTTATAATAATTATTTTCTCCGCGCT 779  
Qy 301 CCTATGTATTTATTTTCAGAAATGAGCCGTACAGCTGCTACAGCAATTACCTACCGGTACC 360  
Db 780 CCTATGTATTTATTTTCAGAAATGAGCCGTACAGCTGCTACAGCAATTACCTACCGGTACC 839  
Qy 361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCCCTGACTGGGCAACAAGTGCATCCAAAT 420  
Db 840 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCCCTGACTGGGCAACAAGTGCATCCAAAT 899  
Qy 421 GACTTGGCGAGTCTTTTGTAGTCTCAGTCTGCTTTGACTGTGTGTACCGCCATCTTT 480  
Db 900 GACTTGGCGAGTCTTTTGTAGTCTCAGTCTGCTTTGACTGTGTGTACCGCCATCTTT 959  
Qy 481 CAATGTCCAGAGTGGCCATCTTTGTGTAGCAACTGTGCCCAAGTGCATGTTGTCCA 540  
Db 960 CAATGTCCAGAGTGGCCATCTTTGTGTAGCAACTGTGCCCAAGTGCATGTTGTCCA 1019  
Qy 541 ACTTGGCGGGGCCCTTTGGGATCCATTGGCACTTGGCTATGGAGAAAGTGGCTAATTCA 600  
Db 1020 ACTTGGCGGGGCCCTTTGGGATCCATTGGCACTTGGCTATGGAGAAAGTGGCTAATTCA 1079  
Qy 601 GTACTTTTCCCTGTAAATATGCTCTCTGGATGTGAATAACTCTGCCACACAGAA 660  
Db 1080 GTACTTTTCCCTGTAAATATGCTCTCTGGATGTGAATAACTCTGCCACACAGAA 1139  
Qy 661 AAAGCAGACCATGAAGAGCTCTGTAGTGTAGGCCCTTATTCCTGTCGCGCCCTGGTCT 720  
Db 1140 AAAGCAGACCATGAAGAGCTCTGTAGTGTAGGCCCTTATTCCTGTCGCGCCCTGGTCT 1199  
Qy 721 TCTGTAAATGGCAAGGCTCTCTGGATGTGTAAATGCCCAATCTGATGCATCAGCATAG 780  
Db 1200 TCTGTAAATGGCAAGGCTCTCTGGATGTGTAAATGCCCAATCTGATGCATCAGCATAG 1259  
Qy 781 TCCATTTACACCTCAGGGAGGATATAGTTTTTCTTGCTACACACATTAATCTTCT 840  
Db 1260 TCCATTTACACCTCAGGGAGGATATAGTTTTTCTTGCTACACACATTAATCTTCT 1319  
Qy 841 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTTTGGCTTTTCACTTCATGTTAGTCTTA 900  
Db 1320 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTTTGGCTTTTCACTTCATGTTAGTCTTA 1379  
Qy 901 GAGAAACAGGAAATACGATGCTACCGAGTCTTTCGCAATCGTACAGCTGATAGA 960  
Db 1380 GAGAAACAGGAAATACGATGCTACCGAGTCTTTCGCAATCGTACAGCTGATAGA 1439  
Qy 961 ACACGCAAGCAAGCTGAAATTTTGTACCGACTTGGCTAAATGGTATAGGCGACGA 1020  
Db 1440 ACACGCAAGCAAGCTGAAATTTTGTACCGACTTGGCTAAATGGTATAGGCGACGA 1499  
Qy 1021 TTGACTTGGGAAGCGACTCTCTGATCTATTTCATGAAGGAATTCGAACAGCCATTATGAAT 1080  
Db 1500 TTGACTTGGGAAGCGACTCTCTGATCTATTTCATGAAGGAATTCGAACAGCCATTATGAAT 1559  
Qy 1081 AGCGACTGTCTAGTCTTTGACACACGATGTCACAGCTTTTTCAGAGAAATGGCAATTTA 1140  
Db 1560 AGCGACTGTCTAGTCTTTGACACACGATGTCACAGCTTTTTCAGAGAAATGGCAATTTA 1619  
Qy 1141 GGCATCAATGTAAGTATTTCCATGTTGAAATGGCAATCAAAATTTTTCGCCAGTGT 1200  
Db 1620 GGCATCAATGTAAGTATTTCCATGTTGAAATGGCAATCAAAATTTTTCGCCAGTGT 1679  
Qy 1201 TTAATACTTCAGTTTCACAGAAATAAGGCACCATCTGTCTGCCAACCTTAAACTCTTT 1260  
Db 1680 TTAATACTTCAGTTTCACAGAAATAAGGCACCATCTGTCTGCCAACCTTAAACTCTTT 1739  
Qy 1261 CGGTAGGTGGAAGC 1274  
Db 1740 CGGTAGGTGGAAGC 1753

; Sequence 25457, Application US/10357930  
; Publication No. US20040259086A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Endege, Wilson  
; APPLICANT: Monahan, John  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
; FILE REFERENCE: MRI-007BCN  
; CURRENT APPLICATION NUMBER: US/10/357,930  
; PRIOR FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: 09/785,276  
; PRIOR FILING DATE: 2003-02-16  
; PRIOR APPLICATION NUMBER: 60/183,319  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: 60/189,862  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/207,454  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/211,314  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/219,007  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/255,281  
; PRIOR FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 62232  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 25457  
; LENGTH: 2924  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1, 57, 2923, 2924  
; OTHER INFORMATION: n = A,T,C or G  
; US-10-357-930-25457

Query Match 100.0%; Score 1274; DB 18; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 3,1e-301;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TTTCTTTAGTGTGTTATGGTCCATTTTCTATTTTACGATTTTATTTCTATGATGCTAT 60  
Db 480 TTTCTTTAGTGTGTTATGGTCCATTTTCTATTTTACGATTTTATTTCTATGATGCTAT 539  
Qy 61 CCAAGACCATTAAGGGAGTTCCACATGTTTTCCGGAACATTTTGAAGAGAGCTTATC 120  
Db 540 CCAAGACCATTAAGGGAGTTCCACATGTTTTCCGGAACATTTTGAAGAGAGCTTATC 599  
Qy 121 CAGGTGACAGATCCCTAATAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT 180  
Db 600 CAGGTGACAGATCCCTAATAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT 659  
Qy 181 TTAATCCTATTTTCTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAACTTTAAAGGACT 240  
Db 660 TTAATCCTATTTTCTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAACTTTAAAGGACT 719  
Qy 241 TATGGCATGTAACATTAATTAAGTAAGTCATGGTTATAATTAATTTTCTCTGCGCT 300  
Db 720 TATGGCATGTAACATTAATTAAGTAAGTCATGGTTATAATTAATTTTCTCTGCGCT 779  
Qy 301 CCTTATGTTATTTTTCAGAAATGAGCCGTACAGCTGCTACAGCATTTACTACCGGTACC 360  
Db 780 CCTTATGTTATTTTTCAGAAATGAGCCGTACAGCTGCTACAGCATTTACTACCGGTACC 839  
Qy 361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCCCTGACTGGCACAACCTGCATCCAAAT 420  
Db 840 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCCCTGACTGGCACAACCTGCATCCAAAT 899  
Qy 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTGACTGTGTGTACCGCCATCTTT 480  
Db 900 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTTGTAGTGTGTGTACCGCCATCTTT 959



481 CAATGTCAGAGTGGCCATCTTTGTTAGCAACTCTGCGCCCAAGCTCACATGTTGTCCA 540  
Db |||||  
960 CAATGTCAGAGTGGCCATCTTTGTTAGCAACTCTGCGCCCAAGCTCACATGTTGTCCA 1019  
Qy 541 ACTTGGCGGGCCCTTTGGATCCATTCGCACTTGGCTATGAGAAAGTGGCTAAATCA 600  
Db |||||  
1020 ACTTGGCGGGCCCTTTGGATCCATTCGCACTTGGCTATGAGAAAGTGGCTAAATCA 1079  
Qy 601 GTACTTTTCCCTGTAAATATATCGTCTCTGATGTGAATAAATCTCTGCCACACACAGAA 660  
Db |||||  
1080 GTACTTTTCCCTGTAAATATATCGTCTCTGATGTGAATAAATCTCTGCCACACACAGAA 1139  
Qy 661 AAAGCAGACCATGAGAGCTCTGAGATTAGGCTTATCTGCTCCGTGCGCTGGTCT 720  
Db |||||  
1140 AAAGCAGACCATGAGAGCTCTGAGATTAGGCTTATCTGCTCCGTGCGCTGGTCT 1199  
Qy 721 TCCTGTAAATGCAAGCTCTCTGATGCTGTAATGCCCATCTGATGCATCAGCATAG 780  
Db |||||  
1200 TCCTGTAAATGCAAGCTCTCTGATGCTGTAATGCCCATCTGATGCATCAGCATAG 1259  
Qy 781 TCATTACAAACCTCAGGAGAGGATATAGTTTTTCTGCTACAGACATTAATCTTCT 840  
Db |||||  
1260 TCATTACAAACCTCAGGAGAGGATATAGTTTTTCTGCTACAGACATTAATCTTCT 1319  
Qy 841 GGTGCTGTTGACTGGGTGATGATGCTCTGTTTGGCTTTCACCTTCATGTTAGTCTTA 900  
Db |||||  
1320 GGTGCTGTTGACTGGGTGATGATGCTCTGTTTGGCTTTCACCTTCATGTTAGTCTTA 1379  
Qy 901 GAGAAACAGGAAATATACATGCTCACCAGAGTCTTCCGCAATCGTACAGCTGATAGGA 960  
Db |||||  
1380 GAGAAACAGGAAATATACATGCTCACCAGAGTCTTCCGCAATCGTACAGCTGATAGGA 1439  
Qy 961 ACACGCAAGCAAGCTGAAATTTTGTATCCGACTTGAAGCTTAATGCTCATGGGACCA 1020  
Db |||||  
1440 ACACGCAAGCAAGCTGAAATTTTGTATCCGACTTGAAGCTTAATGCTCATGGGACCA 1499  
Qy 1021 TTGACTTGGGAAGGCTCTCGATCTATTCATGAAGGAATTCGAACAGCCATTAATGAAT 1080  
Db |||||  
1500 TTGACTTGGGAAGGCTCTCGATCTATTCATGAAGGAATTCGAACAGCCATTAATGAAT 1559  
Qy 1081 AGCAGCTGTCTAGTCTTTGACACAGCAATTCGACAGCTTTTTCGAGAAATGGCAATTTA 1140  
Db |||||  
1560 AGCAGCTGTCTAGTCTTTGACACAGCAATTCGACAGCTTTTTCGAGAAATGGCAATTTA 1619  
Qy 1141 GGCATCAATGTAACTATTTCCATGTTGAAATGGCAATTCGACAGCTTTTTCGCCAGTGT 1200  
Db |||||  
1620 GGCATCAATGTAACTATTTCCATGTTGAAATGGCAATTCGACAGCTTTTTCGCCAGTGT 1679  
Qy 1201 TTAACCTTCAGTTTCACAGAAATAGGCAACCCATCTGCTGCCAACCTTAAACTCTTT 1260  
Db |||||  
1680 TTAACCTTCAGTTTCACAGAAATAGGCAACCCATCTGCTGCCAACCTTAAACTCTTT 1739  
Qy 1261 CGGTAGGTGGAAGC 1274  
Db |||||  
1740 CGGTAGGTGGAAGC 1753

RESULT 4  
US-10-357-930-28822  
; Sequence 28822, Application US/10357930  
; Publication No. US20040259086A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Endege, Wilson  
; APPLICANT: Monahan, John  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
; FILE REFERENCE: MRI-007BCN  
; CURRENT APPLICATION NUMBER: US/10/357,930  
; CURRENT FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: 09/785,276

;; PRIOR FILING DATE: 2003-02-16  
;; PRIOR APPLICATION NUMBER: 60/183,319  
;; PRIOR FILING DATE: 2000-02-17  
;; PRIOR APPLICATION NUMBER: 60/189,862  
;; PRIOR FILING DATE: 2000-03-16  
;; PRIOR APPLICATION NUMBER: 60/207,454  
;; PRIOR FILING DATE: 2000-05-25  
;; PRIOR APPLICATION NUMBER: 60/211,314  
;; PRIOR FILING DATE: 2000-06-09  
;; PRIOR APPLICATION NUMBER: 60/219,007  
;; PRIOR FILING DATE: 2000-07-18  
;; PRIOR APPLICATION NUMBER: 60/255,281  
;; PRIOR FILING DATE: 2000-12-13  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 28822  
;; LENGTH: 2924  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: 1, 57, 2923, 2924  
;; OTHER INFORMATION: n = A,T,C or G  
US-10-357-930-28822

Query Match 100.0%; Score 1274; DB 18; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 3.le-301;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTCTTTAGTTGTTATGCTCCATTTCTATTTTAGCAATTTATTATTCTATGTAGTCTAT 60  
Db |||||  
480 TTTCTTTAGTTGTTATGCTCCATTTCTATTTTAGCAATTTATTATTCTATGTAGTCTAT 539  
Qy 61 CCAAGAGCAGATTAAGGGAGTTCCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 120  
Db |||||  
540 CCAAGAGCAGATTAAGGGAGTTCCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 599  
Qy 121 CAGTGTACAGATCTTAATAAAGTGACATTCAGTGTAATTTTATTTTATATATCTTTT 180  
Db |||||  
600 CAGTGTACAGATCTTAATAAAGTGACATTCAGTGTAATTTTATTTTATATATCTTTT 659  
Qy 181 TTAATCTTATTTTCTCTCTCTTGTCTCAGTAAATTTTGTATGAACTTTTAAAGAGACT 240  
Db |||||  
660 TTAATCTTATTTTCTCTCTCTTGTCTCAGTAAATTTTGTATGAACTTTTAAAGAGACT 719  
Qy 241 TATGCAATGTAACATTAATTTATAAAGTAAGTCAATGTTTATTAATTTTCTCTGCT 300  
Db |||||  
720 TATGCAATGTAACATTAATTTATAAAGTAAGTCAATGTTTATTAATTTTCTCTGCT 779  
Qy 301 CTTATGTATTTATTTTCAAGATGAGCGTCAAGTGTACAGCATTAACCTACCGTACC 360  
Db |||||  
780 CTTATGTATTTATTTTCAAGATGAGCGTCAAGTGTACAGCATTAACCTACCGTACC 839  
Qy 361 TCGAAGTGTCCACCATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db |||||  
840 TCGAAGTGTCCACCATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899  
Qy 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
Db |||||  
900 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959  
Qy 481 CAATGTACAGTGGCCATCTTGTGTAGCAACTGTCGCCCAAGCTCACATGTTGTCCA 540  
Db |||||  
960 CAATGTACAGTGGCCATCTTGTGTAGCAACTGTCGCCCAAGCTCACATGTTGTCCA 1019  
Qy 541 ACTTGGCGGGCCCTTTGGGATCCATTCGCACTTGGCTATGAGAAAGTGGCTAAATCA 600  
Db |||||  
1020 ACTTGGCGGGCCCTTTGGGATCCATTCGCACTTGGCTATGAGAAAGTGGCTAAATCA 1079  
Qy 601 GTACTTTTCCCTGTAAATATAGGCTCTTCTGATGTGAATAAATCTCTGCCACACAGAA 660  
Db |||||  
1080 GTACTTTTCCCTGTAAATATAGGCTCTTCTGATGTGAATAAATCTCTGCCACACAGAA 1139

Qy	661	AAAGCAGACCAATGAAGAGCTCTGTGAGTCTTAGGGCTTATTCTCTGCGTGCCCTGGTGC	720
Db	1140		1199
		AAAGCAGACCAATGAAGAGCTCTGTGAGTCTTAGGGCTTATTCTCTGCGTGCCCTGGTGC	
Qy	721	TCCGTGTAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCAATCAGCATAA	780
Db	1200		1259
		TCCGTGTAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCAATCAGCATAA	
Qy	781	TCCATTACAACCCCTACAGGGAGAGGATATAGTATTTTCTTGCTACAGACATTAATCTTCCT	840
Db	1260		1319
		TCCATTACAACCCCTACAGGGAGAGGATATAGTATTTTCTTGCTACAGACATTAATCTTCCT	
Qy	841	GGTGCTCTTGACTGGGTGATGATGCAGTCCCTGTGTTTGGCTTTCACATCTGTTAGTCTTA	900
Db	1320		1379
		GGTGCTCTTGACTGGGTGATGATGCAGTCCCTGTGTTTGGCTTTCACATCTGTTAGTCTTA	
Qy	901	GAGAAAACAGGAAAATAACGATGGTCAACAGCAGTCTTTCGCAATCGCTACAGCTGATAGGA	960
Db	1380		1439
		GAGAAAACAGGAAAATAACGATGGTCAACAGCAGTCTTTCGCAATCGCTACAGCTGATAGGA	
Qy	961	ACAGCAAGCAAGCTGAAAAATTTTGGCTTACCGACTTGAGCTTAAATGTCATAGGCGACGA	1020
Db	1440		1499
		ACAGCAAGCAAGCTGAAAAATTTTGGCTTACCGACTTGAGCTTAAATGTCATAGGCGACGA	
Qy	1021	TTGACTTTGGGAAGGCGATCCCTCGATCTATTTCATGAAGGAATGCAACAGCCATTATGAAT	1080
Db	1500		1559
		TTGACTTTGGGAAGGCGATCCCTCGATCTATTTCATGAAGGAATGCAACAGCCATTATGAAT	
Qy	1081	AGCGATGCTGTAGTCTTTTGAACACAGCAATTCACAGCTTTTTCGAGAAAAATGGCAATTTA	1140
Db	1560		1619
		AGCGATGCTGTAGTCTTTTGAACACAGCAATTCACAGCTTTTTCGAGAAAAATGGCAATTTA	
Qy	1141	GGCATCAATGTAACTATTTCATGTGTGAAATGGCAATCAACATTTTTCTGGCCAGTGT	1200
Db	1620		1679
		GGCATCAATGTAACTATTTCATGTGTGAAATGGCAATCAACATTTTTCTGGCCAGTGT	
Qy	1201	TTAAAACTTCAGTTTTCAGAAAAATAGGCAACCCATCTGTCTGCCAACCTTAAACCTCTTT	1260
Db	1680		1739
		TTAAAACTTCAGTTTTCAGAAAAATAGGCAACCCATCTGTCTGCCAACCTTAAACCTCTTT	
Qy	1261	CGGTAGGTGGAAGC 1274	
Db	1740		
		CGGTAGGTGGAAGC 1753	

## RESULT 5

```

US-09-764-891-8070
;
; Sequence 8070, Application US/09764891
; Publication No. US20030077808A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;
; FILE REFERENCE: PC006
;
; CURRENT APPLICATION NUMBER: US/09/764,891
;
; CURRENT FILING DATE: 2001-01-17
;
; Prior application data removed - consult PALM or file wrapper
;
; NUMBER OF SEQ ID NOS: 10231
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 8070
;
; LENGTH: 6107
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
US-09-764-891-8070

```

	Query Match	100.0%	Score 1274;	DB 10;	Length 6107;
	Best Local Similarity	100.0%;	Pred. NO. 4,4e-301;		
	Matches 1274;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TTTCTTTAGTGTGTTATATGGTCCATTTTCTATTTTAGCATTTATTTCTATGTAGTCTAT	60		
Db	3164	TTTCTTTAGTGTGTTATATGGTCCATTTTCTATTTTAGCATTTATTTCTATGTAGTCTAT	3223		
Qy	61	CCAAAGACGATTAAGGGAGTTCACATGTTTTTCGGAAACATTTTGAAGAGAGAGCTTATC	120		

[illegible]

Db 4304 GCATCAATGTAACATATTTCCATGTTGGAATGCAATCAACATTTTCTGGCCAGTGT 4363  
Qy 1201 TTAAACATTCAGTTTACAGAAAAATAAGCCATCTGTCTGCCAATCTTAACTCTTT 1260  
Db 4364 TTAAACATTCAGTTTACAGAAAAATAAGCCATCTGTCTGCCAATCTTAACTCTTT 4423  
Qy 1261 CGTAGTGGAGC 1274  
Db 4424 CGTAGTGGAGC 4437

RESULT 6  
US-10-108-260A-269  
; Sequence 269, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108, 260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 269  
; LENGTH: 2829  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-108-260A-269

Query Match 99.9%; Score 1272.4; DB 17; Length 2829;  
Best Local Similarity 99.9%; Pred. No. 7.4e-301;  
Matches 1273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTTCTTTAGTGTGTTATGTCCTATTTTCTATTTTACATTTATTTCTATGTCAT 60  
Db 1134 TTTCTTTAGTGTGTTATGTCCTATTTTCTATTTTACATTTATTTCTATGTCAT 1193

Qy 61 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAAAGAGAGCTTATC 120  
Db 1194 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAAAGAGAGCTTATC 1253

Qy 121 CAGGTACAGATCCTAATAAAGTGCAATTCAGTGTAAATTTTATTTTATATCTTTT 180  
Db 1254 CAGGTACAGATCCTAATAAAGTGCAATTCAGTGTAAATTTTATTTTATATCTTTT 1313

Qy 181 TTAATCCTATTTTCT 240  
Db 1314 TTAATCCTATTTTCT 1373

Qy 241 TATGGCATGTAAACATTTATTTATAAAGTAAGTCAATGTTATTTTCTCTCTCTCT 300  
Db 1374 TATGGCATGTAAACATTTATTTATAAAGTAAGTCAATGTTATTTTCTCTCTCTCT 1433

Qy 301 CTTATGTATTTATTTTACAGAAATAGAGCCGTGACATCTGCTACAGATTTACCTACCGTACC 360  
Db 1434 CTTATGTATTTATTTTACAGAAATAGAGCCGTGACATCTGCTACAGATTTACCTACCGTACC 1493

Qy 361 TCGAAGTGTCCACCATCCAGAGGTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 1494 TCGAAGTGTCCACCATCCAGAGGTCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553

Qy 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
Db 1554 GACTTGGCGAGTCTTTTGTAGTGTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1613

Qy 481 CAATGTACAGTGTGCTCTTTTGTAGTGTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
Db 1614 CAATGTACAGTGTGCTCTTTTGTAGTGTCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1673

Qy 541 ACTTGGCGGGGCGCTTTTGGGATCCATTCGCAACTTGGCTATGGAAGAGTGGCTTAATCA 600  
Db 1674 ACTTGGCGGGGCGCTTTTGGGATCCATTCGCAACTTGGCTATGGAAGAGTGGCTTAATCA 1733

Qy 601 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATAAATACTCTGCCACACACAGAA 660  
Db 1734 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATAAATACTCTGCCACACACAGAA 1793

Qy 661 AAAGCAGACCATGAAGAGCTCTGTAGTGTGAGGCTTATTCCTGTCGTCGCTGCTGCTGCT 720  
Db 1794 AAAGCAGACCATGAAGAGCTCTGTAGTGTGAGGCTTATTCCTGTCGTCGCTGCTGCTGCT 1853

Qy 721 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGTATGCCCATCTGATCAGCAATAAG 780  
Db 1854 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGATCAGCAATAAG 1913

Qy 781 TCCATTACAAACCTTACAGGGAGAGATATAGTGTCTTCTGTCTACAGACATTAATCTTCCT 840  
Db 1914 TCCATTACAAACCTTACAGGGAGAGATATAGTGTCTTCTGTCTACAGACATTAATCTTCCT 1973

Qy 841 GGTGCTGTGACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db 1974 GGTGCTGTGACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2033

Qy 901 GAGAAACAGGAAAAATACGATGGTCAACAGAGTCTTTCGCAATCGTACAGCTGATAGGA 960  
Db 2034 GAGAAACAGGAAAAATACGATGGTCAACAGAGTCTTTCGCAATCGTACAGCTGATAGGA 2093

Qy 961 ACACGCAAGAGCTGAAAAATTTTGTCTTACCGACTTACCGACTTAAATGCTATAGCGACGA 1020  
Db 2094 ACACGCAAGAGCTGAAAAATTTTGTCTTACCGACTTAAATGCTATAGCGACGA 2153

Qy 1021 TTGACTTGGGAAGGACTCTCTCGATCTATTATGAAAGGAATTCGAACAGCCATATGAAAT 1080  
Db 2154 TTGACTTGGGAAGGACTCTCTCGATCTATTATGAAAGGAATTCGAACAGCCATATGAAAT 2213

Qy 1081 AGCGACTCTGTAGTCTTTGACACACAGCATTTGCAAGCTTTTTCGAGAAAAATGGCAATTTA 1140  
Db 2214 AGCGACTCTGTAGTCTTTGACACACAGCATTTGCAAGCTTTTTCGAGAAAAATGGCAATTTA 2273

Qy 1141 GGCAATCAATGTAATTTTCCATGTTGAAATGCAATGCAATCAAACTTTCTGGCCAGTGT 1200  
Db 2274 GGCAATCAATGTAATTTTCCATGTTGAAATGCAATGCAATCAAACTTTCTGGCCAGTGT 2333

Qy 1201 TTAACATTCAGTTTACAGAAAAATAAGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
Db 2334 TTAACATTCAGTTTACAGAAAAATAAGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2393

Qy 1261 CGGTAGTGGAGC 1274  
Db 2394 CGGTAGTGGAGC 2407

RESULT 7  
US-10-641-643-588  
; Sequence 588, Application US/10641643  
; Publication No. US20040077003A1  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; Susan G. Stuart  
; Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL  
; GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641,643  
FILING DATE: 14-Aug-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 588:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2440 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: SPLNFET01  
CLONE: 29244  
SEQUENCE DESCRIPTION: SEQ ID NO: 588 :  
US-10-641-643-588

Query Match 81.9%; Score 1042.8; DB 17; Length 2440;  
Best Local Similarity 99.8%; Pred. No. 1.2e-244;  
Matches 1044; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

229 TTTAAAGGACTATGGCATGTAAACATTTATTAAGTAAGTCAATGCTTATATT 288  
Db  
462 TTTGAAGGACTATGGCATGTAAACATTTATTAAGTAAGTCAATGCTTATATT 521  
Qy  
289 TTTCTCCTGCTTCTTATGTTATTTTTCAGAAATGAGCGCTCAGACTGCTACAGCATTA 348  
Db  
522 TTTCTCCTGCTTCTTATGTTATTTTTCAGAAATGAGCGCTCAGACTGCTACAGCATTA 581  
Qy  
349 CTTACCGGTACTCTGAAGTGTCCACATCCAGAGGGTGCCTGCCCTGACTGGCAAACT 408  
Db  
582 CTTACCGGTACTCTGAAGTGTCCACATCCAGAGGGTGCCTGCCCTGACTGGCAAACT 641  
Qy  
409 GCATCCAAATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTGAATGTGTTA 468  
Db  
642 GCATCCAAATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTGAATGTGTTA 701  
Qy  
469 CGGCCCATCTTCAATGTCCAGAGTGGCCATCTTGTGTAGCAACTGTCCGCCAAAGCTC 528  
Db  
702 CGGCCCATCTTCAATGTCCAGAGTGGCCATCTTGTGTAGCAACTGTCCGCCAAAGCTC 761  
Qy  
529 ACATGTGTCCAACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAA 588  
Db  
762 ACATGTGTCCAACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAA 821  
Qy  
589 GTGGCTAATTCAGTACTTTTCCCTGTAAATATGCTCTTCGGATGTGAATTAATCTGTG 648  
Db  
822 GTGGCTAATTCAGTACTTTTCCCTGTAAATATGCTCTTCGGATGTGAATTAATCTGTG 881  
Qy  
649 CCACACACAGAAAGACAGACCATGAAGAGCTGTGTAGTGTAGGCTTATTCCTGTCCG 708  
Db  
882 CCACACACAGAAAGACAGACCATGAAGAGCTGTGTAGTGTAGGCTTATTCCTGTCCG 941  
Qy  
709 TGCCCTGTGTCTTCTGTAAATGGCAAGGCTCTCTGGATGTGTAAATGCCCATCTGATG 768  
Db  
942 TGCCCTGTGTCTTCTGTAAATGGCAAGGCTCTCTGGATGTGTAAATGCCCATCTGATG 1001  
Qy  
769 CATCAGCATAGTCCATTAACACCCCTACAGGAGAGGATATAGTTTTTCTTCTGCTACAG 828  
Db  
1002 CATCAGCATAGTCCATTAACACCCCTACAGGAGAGGATATAGTTTTTCTTCTGCTACAG 1061  
Qy  
829 ATTAATCTTCTGCTGTCTGTGCTGGGTGATGATGCACTGCTGCTTGTGGCTTCACTTC 888  
Db  
1062 ATTAATCTTCTGCTGTCTGTGCTGGGTGATGATGCACTGCTGCTTGTGGCTTCACTTC 1121

889 ATGTTAGTCTTAGAGAAACAGGAAATAATAGATGTCACAGAGTCTTCGCAATCGTA 948  
Db  
1122 ATGTTAGTCTTAGAGAAACAGGAAATAATAGATGTCACAGAGTCTTCGCAATCGTA 1181  
Qy  
949 CAGCTGATAGGAACACGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGT 1008  
Db  
1182 CAGCTGATAGGAACACGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGT 1241  
Qy  
1009 CATAGCGACGATTTGACTTTGGGAAGCGACTCTCTCGATCTATTTCATGAAGGAATGCAACA 1068  
Db  
1242 CATAGCGACGATTTGACTTTGGGAAGCGACTCTCTCGATCTATTTCATGAAGGAATGCAACA 1301  
Qy  
1069 GCATTAATGAATAGGAGTCTGTCTTGTGACACAGCAGTTCACAGCTTTTTCAGAA 1128  
Db  
1302 GCATTAATGAATAGGAGTCTGTCTTGTGACACAGCAGTTCACAGCTTTTTCAGAA 1361  
Qy  
1129 AATGGCAATTTAGGCATCAATGTAATTTTCATGTTGTAATGGCAATCAACATTT 1188  
Db  
1362 AATGGCAATTTAGGCATCAATGTAATTTTCATGTTGTAATGGCAATCAACATTT 1421  
Qy  
1189 TCTGCCAGTGTTTAAACTTTCAGTTTTCAGAAATAAGGCCACCATCTGTCTGCCAAC 1248  
Db  
1422 TCTGCCAGTGTTTAAACTTTCAGTTTTCAGAAATAAGGCCACCATCTGTCTGCCAAC 1481  
Qy  
1249 CTAAAACTCTTTCGGTAGGTGAAGC 1274  
Db  
1482 CTAAAACTCTTTCGGTAGGTGAAGC 1507

RESULT 8  
US-10-172-118-1117  
; Sequence 1117, Application US/10172118  
; Publication No. US20030224374A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Hongyue  
; APPLICANT: Linsley, Peter  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Chris  
; APPLICANT: Van 't Veer, Laura  
; APPLICANT: Van de Vijver, Marc  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-175-999  
; CURRENT APPLICATION NUMBER: US/10/172,118  
; PRIOR FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/380,770  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 1117  
; LENGTH: 975  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: NM\_005067  
; DATABASE ENTRY DATE: 2001-06-18  
US-10-172-118-1117

Query Match 33.7%; Score 429; DB 17; Length 975;  
Best Local Similarity 73.8%; Pred. No. 1.7e-94;  
Matches 559; Conservative 0; Mismatches 195; Indels 3; Gaps 1;

416 ACAATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTGACTATGTATACGCCCA 475  
Db  
212 ACCAGAGCTGACCTCGCTCTTCGAGTGTCCGCTTGTGACTATGTCTGCTGCTCTTA 271  
Qy  
476 TTTCTCAATGTCCAGTGGCCATCTTGTGTGTAGCAACTGTTCGCCAAAGCTCACAATGT 535  
Db  
272 TTTCTCAGTGTCCAGCGCGGACCTTGTGTGTAACTGCGCCAGAGATTTGAGCTGCT 331  
Qy  
536 GTCCAACTTTCGGGGGCGCTTTGGGATC---CATTCGCAACTTGGCTATGGAGAAAGTGG 592  
Db  
332 GCCCGACGTGCAGGGGGCGCTTCGACGCCCAAGCATCAGGAACTGGCTATGGAGAAAGTGG 391

593 CTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGAGTGAATAACTCTGCCAC 652  
 Db CCTCGGAGTCTGTCTTCCCTGTAAATATGCGCACCACCGGCTGTCTCCCTGACCTGCCAC 451  
 653 ACACAGAAAGACAGACATGAGAGCTCTGTAGCTTTAGGCTTTATTCCTGTCCGTGCC 712  
 Db ATACGGAGAAACAGAAATGAGACATATGAGATACCGTCTCTCTGCCCATGTC 511  
 713 CTGCTGCTTCTGTAAATGAGGAGCTCTCTGAGTGTGTAATGCCCATCTGATGCATC 772  
 Db CTGCTGCTTCTGTAAATGAGGAGCTCTCTGAGTGTGTAATGCCCATCTGATGCATC 571  
 773 AGCAATAGTCCATTAACACCTTACAGGAGAGGATATAGTCTTTCTCTACAGACATTA 832  
 Db CCACAAGAGCATTAACACCTTACAGGAGAGGATATAGTCTTTCTCTACAGACATTA 631  
 833 ATCTTCTGCTGCTGTGACTGCGGTGATGATGACAGTCTCTGTTTGGCTTTTCACTTCATGT 892  
 Db ACTTGCAGGGGCTGTGACTGCGGTGATGATGACAGTCTCTGTTTGGCTTTTCACTTCATGT 691  
 893 TAGTCTTTAGAGAAACAGGAAATACGATGCTCACAGAGTCTTTCGCAATCGTACAGC 952  
 Db TGGTCTGAGAGAAACAGGAAATACGATGCTCACAGAGTCTTTCGCAATCGTACAGC 751  
 953 TGATAGGAACAGCAGCAAGCTGAAATTTTGTACCGACTTACGCTTAATGCTGATATGCTATA 1012  
 Db TCATTGGCCACCGCAAGCGGAGAACTTTGGCTTACAGACTGAGATGAGTGAATGGAAACC 811  
 1013 GCGCAGATTTGCTGGGAGGAGCTCTCTGATCTATTCTATTAAGAGAAATGCAACAGCCA 1072  
 Db GCGGAGATTTGCTGGGAGGAGCTCTCTGATCTATTCTATTAAGAGAAATGCAACAGCCA 871  
 1073 TTATGAATAGCAGCTGTCTGATCTTTTGACACGAGATTCGACAGCTTTTTCGAGAAATG 1132  
 Db TCATGAACAGCAGCTGTCTGATCTTTTGACACGAGATTCGACAGCTTTTTCGAGAAATG 931  
 1133 GCAATTTAGGCAATCAATGTAATTTTCCATGTGTG 1169  
 Db GGAACCTTGAATCAATGTAATTTTCCATGTGTG 968

RESULT 9  
 US-10-342-887-1117  
 ; Sequence 1117, Application US/10342887  
 ; Publication No. US20040058340A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dai, Hongyue  
 ; APPLICANT: He, Yudong  
 ; APPLICANT: Linsley, Peter S.  
 ; APPLICANT: Mao, Mao  
 ; APPLICANT: Roberts, Christopher J.  
 ; APPLICANT: Van 't Veer, Laura Johanna  
 ; APPLICANT: Van de Vijver, Marc J.  
 ; APPLICANT: Bernards, Rene  
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
 ; FILE REFERENCE: 9301-188-999  
 ; CURRENT APPLICATION NUMBER: US/10/342,887  
 ; CURRENT FILING DATE: 2003-01-15  
 ; PRIOR APPLICATION NUMBER: 60/298,918  
 ; PRIOR FILING DATE: 2001-06-18  
 ; PRIOR APPLICATION NUMBER: 60/380,710  
 ; PRIOR FILING DATE: 2002-05-14  
 ; PRIOR APPLICATION NUMBER: 10/172,118  
 ; PRIOR FILING DATE: 2002-06-14  
 ; NUMBER OF SEQ ID NOS: 2699  
 ; SEQ ID NO 1117  
 ; LENGTH: 975  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-342-887-1117  
 Query Match 33.7%; Score 429; DB 17; Length 975;

Best Local Similarity 73.8%; Pred. No. 1.7e-94;  
 Matches 559; Conservative 0; Mismatches 195; Indels 3; Gaps 1;  
 416 ACATGACTTGGGAGTCTTTTGGAGTGTCCAGTCTGCTTTGACTATGTGTACCGGCCA 475  
 Db ACCACGAGCTGACCTCGCTCTTGGAGTGTCCGCTCTGCTTTGACTATGTCTCTGCTCTA 271  
 476 TTCTTCAATGTCAGAGTGGCCATCTTGTGTTAGCAAACTGTGCCCCAAAGCTCACATGTT 535  
 Db TTCTGAGTGGCCAGCGCGGACCTGTGTGTAAACCAATGCCGACAGAGTTGAGCTGCT 331  
 536 GTCCAACTGCGCGGCGCTTTTGGGATC---CATTCGCAACTTGGCTATGGAGAAAGTGG 592  
 Db GCGCGAGCTGTCAGGCGGCGCTTCAAGCGCCAGCATCAGGAACCTGGCTATGAGAAAGTGG 391  
 593 CTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATAAATCTCTGCCAC 652  
 Db CCTCGGAGTCTCTGTTTCCCTGTAAATATGCGCACAAGGCTGTTCCTGACCTGCCAC 451  
 653 ACACAGAAAGACAGACCATGAAGAGCTCTGTGAGTTTGGCCCTTATTCCTGTCGCTGCC 712  
 Db ATACGGAGAAACAGAAATACGATGATGATGATGATGATGATGATGATGATGATGATGAT 511  
 713 CTGCTGCTTCTGTAATGCAAGGCTCTCTGATGCTGTGATGCTGTGATGCTGTGATGCT 772  
 Db CTGCTGCTTCTGCAAGTGGCAGGCGTCCCTGGAAGCTGTGATGCTGTGATGCTGTGATGCT 571  
 773 AGCATAGTCCATTAACACCTTACAGGAGGAGGATATAGTATTTTCTGCTACAGACATTA 832  
 Db CCACAGAGATTTACACCTTACAGGAGGAGGATATAGTATTTTCTGCTACAGACATTA 631  
 833 ATCTTCTGCTGTGTTGACTGCGGTGATGATGCTGCTGTTTGGCTTTTCACTTCATGT 892  
 Db ACTTGCAGGGGCTGTCCACTGCGGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691  
 893 TAGTCTTACAGAAACAGGAAATACGATGCTCACAGAGTCTTTCGCAATCGTACAGC 952  
 Db TGGTCTGAGAAACAGGAAATACGATGCTCACAGAGTCTTTCGCAATCGTACAGC 751  
 953 TGATAGGAACAGCAGCAAGCTGAAATTTTGTACCGACTTACGCTTAATGCTGATATGCTATA 1012  
 Db TCATTGGCCACCGCAAGCGGAGAACTTTGGCTTACAGACTGAGTGAATGGAAACC 811  
 1013 GCGCAGATTTGCTGGGAGGAGCTCTCTGATCTATTCTATTAAGAGAAATGCAACAGCCA 1072  
 Db GCGGAGATTTGCTGGGAGGAGCTCTCTGATCTATTCTATTAAGAGAAATGCAACAGCCA 871  
 1073 TTATGAATAGCAGCTGTCTGATCTTTTGACACGAGATTCGACAGCTTTTTCGAGAAATG 1132  
 Db TCATGAACAGCAGCTGTCTGATCTTTTGACACGAGATTCGACAGCTTTTTCGAGAAATG 931  
 1133 GCAATTTAGGCAATCAATGTAATTTTCCATGTGTG 1169  
 Db GGAACCTTGAATCAATGTAATTTTCCATGTGTG 968

RESULT 10  
 US-10-177-293-426  
 ; Sequence 426, Application US/10177293  
 ; Publication No. US20030124128A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lillie, James  
 ; APPLICANT: Glatt, Karen  
 ; APPLICANT: Zhao, Xumei  
 ; APPLICANT: Gannavarpu, Manjula  
 ; APPLICANT: Kamatkar, Shubhangi  
 ; APPLICANT: Mertens, Maureen  
 ; APPLICANT: Myer, Vic  
 ; APPLICANT: Wang, Youzhen  
 ; APPLICANT: Xu, Yongyao  
 ; APPLICANT: Hoersch, Sebastian  
 ; APPLICANT: Monahan, John  
 ; APPLICANT: Meyers, Rachel E.

; APPLICANT: Bast Jr., Robert C.  
; APPLICANT: Hortobagyi, Gabriel N.  
; APPLICANT: Puzstai, Lajos  
; APPLICANT: Meric, Aysegul  
; APPLICANT: Sahin, Aysegul  
; APPLICANT: Mills, Gordon B.  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,  
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-038  
; CURRENT APPLICATION NUMBER: US/10/177,293  
; PRIOR FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: US 60/299,887  
; PRIOR FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: US 60/301,572  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: US 60/306,501  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US 60/325,002  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/362,585  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 506  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 426  
; LENGTH: 2502  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-177-293-426

Query Match 33.7%; Score 429; DB 15; Length 2502;  
Best Local Similarity 73.8%; Pred. No. 2.7e-94;  
Matches 559; Conservative 0; Mismatches 195; Indels 3; Gaps 1;  
  
QY 416 ACAATGACTGGCGAGCTCTTTTGTAGCAACTGTGCGCCCAAGAGCTCAGATGTT 475  
DB 738 ACCAGAGCTGACCTCGCTCTTCGAGTGTCGGCTGTGTTGACTATGCTTCCCTCCTA 797  
QY 476 TTCTTCAATGTCAGAGTGCGCACTCTTTGTAGCAACTGTGCGCCCAAGAGCTCAGATGTT 535  
DB 798 TTCTGAGTGCCAGGCGCGGCACTGTGTGTAACTGCGCCAGAGTTGAGCTGCT 857  
QY 536 GTCCAACTTGGCGGCGCTTTGGGATC---CATTCGCAACTGTTGGCTATGGAAGAGTGG 592  
DB 858 GCGCGAGTGCGAGGCGCGCTTACGCCAGCATCAGGAACCTGGCTATGGAAGAGTGG 917  
QY 593 CTAAATTCAGTACTTTTCCCTCTGAATATGCGTCTTCTGATGTGAATACTCTGCCAC 652  
DB 918 CTTGCGCATGCTCTTCTTCCCTGTAAATGATGCGCAACCGGCTGTTCCTGACCTGCACC 977  
QY 653 ACACAGAAAAAGCAGACCATGAAGAGCTGTGAGTGTAGGCTTATCTGTGCTGCGTGC 712  
DB 978 ATACGAGAGAACAGAGCATGATGATATGATGATGATGATGATGATGATGATGATGATG 1037  
QY 713 CTGGTCTTCTGTAATGGCAAGGCTCTCTGATGCTGTAAATGCGCCCATCTGATGATC 772  
DB 1038 CTGGTCTTCTGTAATGGCAAGGCTCTCTGATGCTGTAAATGCGCCCATCTGATGATC 1097  
QY 773 AGCATAGTCCATTACAACTACAGGAGAGATAGTATTTCTGTACAGACATTA 832  
DB 1098 CCCAAGAGCATTTACCCCTTACAGGAGAGATGCTCTTCTGATGATGATGATGATGAT 1157  
QY 833 ATCTTCTGCTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 892  
DB 1158 ACTTGCAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1217  
QY 893 TAGTCTTACAGAAACAGGAAAAATACGATGGTACAGAGAGTCTTTCGCAATGTCACAG 952  
DB 1218 TGGTGTGAGAGAAACAGAGAGAGTACGAAGGCGCCAGCAGATTTTGTGCAATGCTCTG 1277  
QY 953 TGATAGGAACAGCAAGAGCTGAAAAATTTTGTACCGACTTGTAGCTTAATGTCATA 1012

DB 1278 TCATTGGCAGCCCGCAAGCCGAGAACTTTGCTACAGACTGGAGTTGAATGGGAACC 1337  
QY 1013 GCGCAGCATTTGACTTGGGAAGCGACTCTCTGATCTATTTCATGAAGGAATTCGCAACAGCCA 1072  
DB 1338 GCGCGAGATTGACTTGGGAGGCCAGCCCGCTTTCGATTTCATGACGGTGTGCTGCGGCCA 1397  
QY 1073 TTATGAATAGCGACTGTCTAGTCTTTTGACACACGCAATTCGACAGCTTTTTCGAGAAAATG 1132  
DB 1398 TCATGAACAGCGACTGCTCTTGTCTTTCGACACAGCCATAGCACATCTTTTTCGAGATAATG 1457  
QY 1133 GCAATTTAGCATCAATGTAACCTATTTCATGATGTTG 1169  
DB 1458 GGAACCTTGAATCAATGTTACTATTTCATGATGTTG 1494  
  
RESULT 11  
US-10-177-293-424  
; Sequence 424, Application US/10177293  
; Publication No. US20030124128A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Gannavarpu, Manjula  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Mertens, Maureen  
; APPLICANT: Myer, Vic  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Monahan, John  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Bast Jr., Robert C.  
; APPLICANT: Hortobagyi, Gabriel N.  
; APPLICANT: Puzstai, Lajos  
; APPLICANT: Meric, Aysegul  
; APPLICANT: Sahin, Aysegul  
; APPLICANT: Mills, Gordon B.  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,  
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-038  
; CURRENT APPLICATION NUMBER: US/10/177,293  
; CURRENT FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: US 60/299,887  
; PRIOR FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: US 60/301,572  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: US 60/306,501  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US 60/325,002  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/362,585  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 506  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 424  
; LENGTH: 2240  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-177-293-424

Query Match 33.5%; Score 427.4; DB 15; Length 2240;  
Best Local Similarity 73.7%; Pred. No. 6.2e-94;  
Matches 558; Conservative 0; Mismatches 196; Indels 3; Gaps 1;  
  
QY 416 ACAATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTCTGCTTTTGTGATGTTACCGCCCA 475  
DB 738 ACCAGAGCTGACCTCGCTCTTCGAGTGTCCGCTCTGCTTTGCTGATGCTCTGCTCCTA 797  
QY 476 TTCTTCAATGTCAGAGTGCGCACTCTTGTGTAGCAACTGTGCGCCCAAGAGCTCAGATGTT 535



Db	798	TTCTGCAGTCCGAGCGCGGCACCTGGTGTGTAAACCAATGCCGCCAGAAAGTTGAGCTGCT	857
Qy	536	GTCCACCTTCGCGGGGCCCTTTCCGGATC---CATTCGCAACTTGGCTATGGAGAAAGTGG	592
Db	858	GCCCGAGTGCAGGGGGGCCCTGACGCCAGCAATCAGGAACCTGGCTATGGAGAAAGTGG	917
Qy	593	CTAATTCAAGTACTTTTCCCTGTAAATATACGGTCTTCTGGATGTGAAATAACTCTGCCAC	652
Db	918	CCTCGGCAGTCTGTGTTCCTGTAAATATGCCACACGGGCTGTTCCTCTGACCCCTGCACC	977
Qy	653	ACACAGAAAAAGCAGACATGAGAGAGCTCTGTGAGTTTAGCGCTTATTCCTGTCCGTGCC	712
Db	978	ATACGGAGAAACCGAACAATAGGAAGCATATGTGAATACCCGCCCTACTCTCGGCCATGTC	1037
Qy	713	CTGGTGCTTCTCTTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCCATCTCATGCAATC	772
Db	1038	CTGGTGTCTCTGCAAGTGGCAGGGTCCCTGGAAAGCTGTGATGTGCCATCTCATGCACG	1097
Qy	773	AGCATAGTCCATTACAAACCTACAGGAGAGGATATAGTTTTCTGTCTACAGACATTA	832
Db	1098	CCCAACAGACATTACCAACCCCTCAGAAAGAACATCGCTCTTTCTAGCTACAGACATTA	1157
Qy	833	ATCTTCTGTGGTGTGTTGACTGGGGTGATGATGCAAGTCTCTGTTTTTGGCTTTCACTTCATGT	892
Db	1158	ACTTGCACGGGGCTGTGCAGTGGGTGATGTCAGTCAATGTTTTTGGCCATCACTTCATGC	1217
Qy	893	TAGTCTTAGAGAAAACAGAAAAATACAGATGGTACACGAGAGTCTTCGCAATGTGTACGC	952
Db	1218	TGGTGTGGAGAAAAACAGAGAAGTACGAAGGCCACCAAGCAGTTTTTTTGGCATGTCTCTGC	1277
Qy	953	TGATAGGAACACGCAAGCAAGCTGAAAATTTTGCTTACCGACTTGAGCTAAATGGTCATA	1012
Db	1278	TCATTGGCACCCGACAGCAACCCGAGAACTTTTGGCTACAGACTGGAGTTGAAATGGGAACC	1337
Qy	1013	GGCGACGATTGACTTGGGAAGCGACTCTCTCGATCTATTCTATGAAAGGAATTTGCAACAGCCA	1072
Db	1338	GGCGGAGATTGACCTCTGGGAGGCCACGCCCTCGATTCTCATGACGGTGTGGCTGCGGCCCA	1397
Qy	1073	TTATGAATAGCGACTGTCTAGTCTTTGACACACGACATTTGCACACCTTTTTTGCAGAAAATG	1132
Db	1398	TCATGAACAGCGACTGCGCTTGTTTTTGCACACGACCATAGCACATCTTTTTTGCAGATATG	1457
Qy	1133	GCAATTTTAGGCATCAATGTAACATAATTTCCATGTGTG	1169
Db	1458	GGAACTTTGGAAATCAATGTGTACTATTTCTACATGTTG	1494

```

RESULT 12
US-10-717-597-137
; Sequence 137, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dorner, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 137
; LENGTH: 2240
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-10-717-597-137

	Query Match Best Local Similarity Matches	33.5%; 73.7%; 558;	Score 427.4; Pred. No. 6.2e-94; Conservative 0;	DB 18; Indels 3; Gaps 1;	Length 2240;
QY	416	ACAATGACTGGCGAGTCTTTTGAGTGTCCAGTCGTCTTGA	CTCTGCTTTGACGTATGTGTACGCCCA	475	
Db	738	ACACGAGCTGACCTCGCTCTTCGAGTGC	CGGTCTGCTTTGACTATGTCTCGCTCCTTA	797	
QY	476	TTCCTCAATGTCAAGATGGGCATCTTTGTTGTAGCAACTGT	CGCCCAAGACTCACATGTT	535	
Db	798	TTCTGCAGTCCGAGCGCGGACCTCGTGTGTAAACAAATGC	CGCCAGAAGTTGAGCTGCT	857	
QY	536	GTTCCAACTTCCCGGGGCCCTTTGGGATC---	CATTCCGCAACTTGGCTATGGAGAAAGTGG	592	
Db	858	GCCCAGCTGCGAGGGCGCCCTGACGCCAGCATCAGGAAC	CTGGCTATGGAGAAGGTGG	917	
QY	593	CTAATTCAGTACTTTTCCCTGTAAATATGCTGTC	TCTGGATGTGAAATAACTCTGCCAC	652	
Db	918	CCTCGGAGTCTCTGTTTTCCCTGTAAATATGCTGAC	CGGCGTCTCCCTGACCCTGCACC	977	
QY	653	ACACAGAAAAGCAGACCATGAAGAGCTCTGTCAGATT	TAGGCCCTTAATCTCTGCTCCGTGCC	712	
Db	978	ATACGGAGAAACCAAGAACATGAAGACATATGTCAA	TACCGCCCCCTACTCTGCCCATGTC	1037	
QY	713	CTGTGCTCTCTGTAAATGGCAAGGCTCTCTGATGCTGT	AATGTGCCCCCATCTGATGCATC	772	
Db	1038	CTGTGCTCTCTGCAAGTGGCAGGGGTCCCTGGAAGCT	GTGATGCCCATCTCATGCACG	1097	
QY	773	AGCATAAGTCCAATACAAACCTCACGGAGAGGATATAG	TTTTTCTTGCTPACAGACATTA	832	
Db	1098	CCACAGAGCAATACCAACCTTCAGSAGAGACATCGT	CTTCTTAGCTACAGACATTA	1157	
QY	833	ATCTCTCTGTGTGTTGACTGGGTGATGATGAGTCTGT	CTGCTTTGGCTTTCACTTCATGT	892	
Db	1158	ACTTGCCAGGGGCTGTGCTGCTGGGTGATGATGAGT	CATGTTTTTGGCCATCCTTTCATGC	1217	
QY	893	TAGTCTTTAGAGAAACAGAAAAATACCATGGT	CACACAGAGTTCTTTCGCAATCGTACAGC	952	
Db	1218	TGTTGCTCGAGAAACAAGAGNATACGAAGCCAC	AGCAGTTTTTTTTGGCCATCGTCTCTGC	1277	
QY	953	TGATAGMAACACGCAAGCTGAAATTTTGCTTACCG	ACTTGAGCTAGCTAAATGGTTCATA	1012	
Db	1278	TCATTTGGCACCCGACGACGCCGAGAAC-TTTG	CTACAGCTGGAGTTGATTTGGGNACC	1337	
QY	1013	GSCGACCATTCACTTGGSAAGCGACTCTCTCGATCT	ATTTCATGTAAGGAATTTGCAACAGCCA	1072	
Db	1338	GCGGAGATTACCTGGSGAGCCACGCCCCGCTCG	ATTCATGACGGTGTGGCTGCGGCCCA	1397	
QY	1073	TTATGAAATAGGACTGTCTAGTCTTTTGCAACAC	CAGCATTCGACAGCTTTTTTGCAGAAAATG	1132	
Db	1398	TCATGAACAGGAGCTGCGCTTGTTTTGCACAC	GAGCCATGACATCTTTTTTGCAGTAATG	1457	
QY	1133	GCAATTTAGGCATCAATGTAACATATTTCCATGT	GTG 1169		
Db	1458	GTAACTTGGAAATCAATGTATCTATTTCTACAT	GTG 1494		

RESULT 13  
US-10-755-889-123  
; Sequence 123, Application US/10755889  
; Publication No. US20040171823A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB  
; TITLE OF INVENTION: PATHWAY  
; FILE REFERENCE: D0284 NP  
; CURRENT APPLICATION NUMBER: US/10755,889  
; CURRENT FILING DATE: 2004-01-13  
; PRIOR APPLICATION NUMBER: U.S. 60/440,068  
; PRIOR FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: U.S. 60/469,757



; PRIOR FILING DATE: 2003-05-12  
; NUMBER OF SEQ ID NOS: 823  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 123  
; LENGTH: 2240  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-755-889-123

Query Match 33.5%; Score 427.4; DB 18; Length 2240;  
Best Local Similarity 73.7%; Pred. No. 6.2e-94;  
Matches 558; Conservative 0; Mismatches 196; Indels 3; Gaps 1;  
QY 416 ACATGACTGGGAGTCTTTTGAAGTGCAGTCTGCTTTGACTATGTTACGCCCA 475  
Db |||||  
-738 ACCACAGAGTACCTCGCTCTTCGAGTGTCCGCTCTGTTGACTATGTTACGCCCTTA 797  
QY 476 TTCTTCAATGTGAGTGGCCATCTGTTTGTAGCAACTGTGCGCCCAAGCTCACATGTT 535  
Db |||||  
798 TTCTGAGTGCACGCGCGGCACCTGTGTGTAAACATGCGCGCAGAGTTGAGCTGCT 857  
QY 536 GTCCAACTTGGCGGCGCCCTTTGGGATC---CATTCGCAACTTGGCTATGGAGAAAGTGG 592  
Db |||||  
858 GCCCGAGTGCAGGGCGCCCTGACCCGAGCATCAGGAACCTGGCTATGGAGAGTGG 917  
QY 593 CTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGATGTGAAATTAACCTGCGCAC 652  
Db |||||  
918 CCTCGCAGTCTGTTTCCCTGTAAATATGCGCACCACCGGCTGTTCCCTGACCCCTGCACC 977  
QY 653 ACACAGAAAACGACCATGAGAGCTCTGTCAGTTTGGGCTTATCTCTGCTCGTGGC 712  
Db |||||  
978 ATACGAGAAAACGAGAAATGAGACATATGTGAATACCGCCCTACTCTCGCCCATGTC 1037  
QY 713 CTGGTCTTCTGTAATGGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATGCATC 772  
Db |||||  
1038 CTGGTCTTCTGCAAGTGCAGGGTCTCTGAGCTGTGATGTCCTATCTCTGTCGAGC 1097  
QY 773 AGCATAAGTCCATTACACCCCTACAGGAGAGATATAGTTTCTTGTACAGACATTA 832  
Db |||||  
1098 CCACAGAGCATTACACCCCTTACAGAAAGACATCGTCTTCTAGCTACAGACATTA 1157  
QY 833 ATCTCTCTGCTGTTGATGCTGGTGATGATGATGATGATGATGATGATGATGATGATG 892  
Db |||||  
1158 ACTTGCAGGGGCTGTCGATGCTGGTGATGATGATGATGATGATGATGATGATGATGATG 1217  
QY 893 TAGTCTTAGAGAAACAGGAAATACATGCTGTCAGCAGAGTCTTCGCAATCGTACAGC 952  
Db |||||  
1218 TGGTCTGAGAAACAGAGAGTACGAGGCGCACAGCAGTTTGTGCAATCGTCTGCTG 1277  
QY 953 TCATAGGAACACGCAAGCTGAAATTTTGGCTTACCGACTTGAGCTGATGATGATGATGATG 1012  
Db |||||  
1278 TCATTTGGCACCAGCAAGCAGGAGCTTTTGGCTTACAGACTGGAGTTGATGAGGAAACC 1337  
QY 1013 GCGGAGATTGACTTGGGAGCGACTCTCGATCTATTCATGAGGAATTCAGACGCA 1072  
Db |||||  
1338 GCGGAGATTGACTTGGGAGCGACCGCCCTCGATTCATGACGCTGCTGCTGCGGCA 1397  
QY 1073 TTATGAATAGGAGCTGCTAGTCTTTGACACAGCAATTCACAGCTTTTTCAGAAATG 1132  
Db |||||  
1398 TCATGACAGCACTGCTTGTGTTTCGACAGCCATACATCTTTTTCAGATATG 1457  
QY 1133 GCAATTTAGGCATCAATGTAATCTTATTCATGTTG 1169  
Db |||||  
1458 GGAACCTTGAATCAATGTTACTATTCTACATGTTG 1494

RESULT 14  
US-09-925-297-84  
; Sequence 84, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 84  
; LENGTH: 1535  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-297-84

Query Match 29.7%; Score 379; DB 9; Length 1535;  
Best Local Similarity 99.7%; Pred. No. 3.7e-82;  
Matches 390; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 885 CTTTCATGTTAGTCTTAGAGAAACAGGAAATAATAC-GATGGTCACCGAGCTTCTTCGCAA 943  
Db 1 CTTTCATGTTAGTCTTAGAGAAACAGGAAATAATACGATGGTCACCGAGCTTCTTCGCAA 60  
QY 944 TCGTACAGCTGATAGACACGACGACGAGCTGAAATTTTCTTACCGACTTGAGCTAA 1003  
Db |||||  
61 TCGTACAGCTGATAGACACGACGACGAGCTGAAATTTTCTTACCGACTTGAGCTAA 120  
QY 1004 ATGGTCAATGAGGAGCATTTGACTTGGGAGCGACTCTCGATCTATTATGAAGGAATTG 1063  
Db |||||  
121 ATGGTCAATGAGGAGCATTTGACTTGGGAGCGACTCTCGATCTATTATGAAGGAATTG 180  
QY 1064 CAACAGCCATTAATGATAGCAGCTGTCTAGTCTTTTGACACCGAGCATTTGACAGCTTTTGG 1123  
Db |||||  
181 CAACAGCCATTAATGATAGCAGCTGTCTAGTCTTTTGACACCGAGCATTTGACAGCTTTTGG 240  
QY 1124 CAGAAATGGCAATTTAGGCATCAATGTAATTTTCCATGTTTGAATGCAATCAAA 1183  
Db |||||  
241 CAGAAATGGCAATTTAGGCATCAATGTAATTTTCCATGTTTGAATGCAATCAAA 300  
QY 1184 CATTTTCTGGCCAGTGTTTAAACCTTCAGTTTTCAGAGAAATAAGGCACCCCATCTGCTG 1243  
Db |||||  
301 CATTTTCTGGCCAGTGTTTAAACCTTCAGTTTTCAGAGAAATAAGGCACCCCATCTGCTG 360  
QY 1244 CCAACTAAACTCTTTTCGGTAGGTGGAAGC 1274  
Db |||||  
361 CCAACTAAACTCTTTTCGGTAGGTGGAAGC 391

RESULT 15  
US-09-960-352-10011  
; Sequence 10011, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; NUMBER OF SEQ ID NOS: 2001-09-24  
; SEQ ID NO 10011  
; LENGTH: 360  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (312)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 43-LIB3058-032-Q1-K1-C4  
US-09-960-352-10011

Mon Apr 25 11:06:39 2005

```
Query Match      23.1%; Score 293.8; DB 9; Length 360;
Best Local Similarity 91.5%; Pred. No. 1.3e-61;
Matches 324; Conservative 0; Mismatches 23; Indels 7; Gaps 1;

QY 910 GAAAAATACGATGGTCACACAGAGTTCTTCGCAATCGTACAGCTGATAGGAACACGCAAG 969
Db 1 GAAAAATATGATGGTCATCAGCAATTCCTCGCAATTTGACAGCTGATAGGAACACGCAAG 60

QY 970 CAAGCTGAAAAATTTTGCTTTACCGACTTGAGCTAAATGGTCAATAGCGGCGGCAATGACTTTGG 1029
Db 61 CAAGCTGAAAAATTTTGCTTTATAGACTTGAGCTAAATGGTCAATAGCGGCGGCAATGACTTTGG 120

QY 1030 GAAGCGACTCCCTCGATCTATTGATGAAGGAATTGCAACAGCCATTATGAATAGCGACTGT 1089
Db 121 GAAGCCACTCCCTCGCTCTATTGATGAGGGAATTGCAACAGCCATTATGAATAGTCACTGC 180

QY 1090 CTAGTCTTTTGACACCAAGCATTGCACAGCTTTTTCAGAAAAATGGCAATTTAGGCAATCAAT 1149
Db 181 CTAGTCTTTTGACACCAAGCATTGCACAGCTCTTTTGCAAAAAATGGCAATTTAGGCAATCAAT 240

QY 1150 GTAACTATTTCATGTTGAAATGGCAATCAAAATTTTCTGGCCAGTGTAAAAAC-- 1207
Db 241 GTAACTATTTCATGTTGAAATGGCAATCAGACATTTTCTGGCCAGTGTAAAAACCA 300

QY 1208 -----TTCAGTTTCAGAAAAATAGGCAACCCATCTGCTGCCAACCTAAAACT 1256
Db 301 TTGCATTCAATNTCACAGAGAATAGGCAACCTGTCTGCTTACCAACCAAAAACT 354
```

Search completed: April 25, 2005, 00:49:28  
Job time : 826 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 19:49:27 ; Search time 5406 Seconds  
(without alignments)

9620.525 Million cell updates/sec

Title: US-10-679-246-1

Perfect score: 1274

Sequence: 1 ttcttttggtgtttatggt.....ctctttcgtagtggaagc 1274

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:\*

- 1: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pna/US099G COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pna/US099G COMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pna/US099H COMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pna/US099I COMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pna/US099J COMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pna/US099K COMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pna/US099L COMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pna/US099M COMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pna/US099N COMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pna/US099O COMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pna/US099P COMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pna/US099Q COMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pna/US099R COMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pna/US099S COMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pna/US099T COMB.seq:\*
- 23: /cgn2\_6/ptodata/1/pna/US099U COMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pna/US099V COMB.seq:\*
- 25: /cgn2\_6/ptodata/1/pna/US099W COMB.seq:\*
- 26: /cgn2\_6/ptodata/1/pna/US099X COMB.seq:\*
- 27: /cgn2\_6/ptodata/1/pna/US099Y COMB.seq:\*
- 28: /cgn2\_6/ptodata/1/pna/US099Z COMB.seq:\*
- 29: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*
- 30: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*
- 31: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*
- 32: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*
- 33: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*
- 34: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*
- 35: /cgn2\_6/ptodata/1/pna/US099G COMB.seq:\*
- 36: /cgn2\_6/ptodata/1/pna/US099H COMB.seq:\*
- 37: /cgn2\_6/ptodata/1/pna/US099I COMB.seq:\*
- 38: /cgn2\_6/ptodata/1/pna/US099J COMB.seq:\*
- 39: /cgn2\_6/ptodata/1/pna/US099K COMB.seq:\*
- 40: /cgn2\_6/ptodata/1/pna/US099L COMB.seq:\*
- 41: /cgn2\_6/ptodata/1/pna/US099M COMB.seq:\*
- 42: /cgn2\_6/ptodata/1/pna/US099N COMB.seq:\*
- 43: /cgn2\_6/ptodata/1/pna/US099O COMB.seq:\*

44: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

45: /cgn2\_6/ptodata/1/pna/US099G COMB.seq:\*

46: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

47: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

48: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

49: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

50: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

51: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

52: /cgn2\_6/ptodata/1/pna/US099G COMB.seq:\*

53: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

54: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

55: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

56: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

57: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

58: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

59: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

60: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

61: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

62: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

63: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

64: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

65: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

66: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

67: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

68: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

69: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

70: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

71: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

72: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

73: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

74: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

75: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

76: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

77: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

78: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

79: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

80: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

81: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

82: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

83: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

84: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

85: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

86: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

87: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

88: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

89: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

90: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

91: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

92: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

93: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

94: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

95: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

96: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

97: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

98: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

99: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

100: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

101: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

102: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

103: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

104: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

105: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

106: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

107: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

108: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

109: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

110: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

111: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

112: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

113: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

114: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

115: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

116: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

```
117: /cgn2_6/ptodata1/pna/US6047_COMB.seq.*
118: /cgn2_6/ptodata1/pna/US6048_COMB.seq.*
119: /cgn2_6/ptodata1/pna/US6049_COMB.seq.*
120: /cgn2_6/ptodata1/pna/US6050_COMB.seq.*
121: /cgn2_6/ptodata1/pna/US6051_COMB.seq.*
122: /cgn2_6/ptodata1/pna/US6052_COMB.seq.*
123: /cgn2_6/ptodata1/pna/US6053_COMB.seq.*
124: /cgn2_6/ptodata1/pna/US6054_COMB.seq.*
125: /cgn2_6/ptodata1/pna/US6055_COMB.seq.*
126: /cgn2_6/ptodata1/pna/US6056_COMB.seq.*
127: /cgn2_6/ptodata1/pna/US6057_COMB.seq.*
128: /cgn2_6/ptodata1/pna/US6058_COMB.seq.*
129: /cgn2_6/ptodata1/pna/US6059_COMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1274	100.0	1274	20	US-09-330-517-1	Sequence 1, Appli
2	1274	100.0	1274	58	US-10-679-246-1	Sequence 1, Appli
3	1274	100.0	1435	21	US-09-359-922-5961	Sequence 5961, Ap
4	1274	100.0	1435	21	US-09-359-922-5961	Sequence 5961, Ap
5	1274	100.0	1435	39	US-09-919-002-5961	Sequence 5961, Ap
6	1274	100.0	1735	29	US-09-644-867-5326	Sequence 5326, Ap
7	1274	100.0	1735	29	US-09-652-918-7745	Sequence 7745, Ap
8	1274	100.0	2270	49	US-10-170-235-39760	Sequence 39760, A
9	1274	100.0	2599	102	US-60-324-185-25086	Sequence 25086, A
10	1274	100.0	2921	29	US-09-644-867-7203	Sequence 7203, Ap
11	1274	100.0	2921	31	US-09-698-013-6733	Sequence 6733, Ap
12	1274	100.0	2921	33	US-09-726-811-5469	Sequence 5469, Ap
13	1274	100.0	2924	34	US-09-785-276A-22959	Sequence 22959, A
14	1274	100.0	2924	34	US-09-785-276A-25457	Sequence 25457, A
15	1274	100.0	2924	34	US-09-785-276A-28822	Sequence 28822, A
16	1274	100.0	2924	53	US-10-357-930-22959	Sequence 22959, A
17	1274	100.0	2924	53	US-10-357-930-25457	Sequence 25457, A
18	1274	100.0	2924	53	US-10-357-930-28822	Sequence 28822, A
19	1274	100.0	6107	1	PCT-US01-01329-2898	Sequence 2898, Ap
20	1274	100.0	6107	1	PCT-US01-01339-8070	Sequence 8070, Ap
21	1274	100.0	6107	34	US-09-764-891-8070	Sequence 8070, Ap
22	1274	100.0	6107	47	US-10-080-090-2898	Sequence 2898, Ap
23	1274	100.0	6107	50	US-10-205-303-2898	Sequence 2898, Ap
24	1272.8	99.9	2270	114	US-60-443-566-1446	Sequence 1446, Ap
25	1272.8	99.9	2270	114	US-60-449-629-24	Sequence 24, Appl
26	1272.8	99.9	2270	115	US-60-455-444-558	Sequence 558, App
27	1272.8	99.9	2270	116	US-60-465-241-558	Sequence 558, App
28	1272.8	99.9	27755	65	US-10-918-754-16927	Sequence 16927, A
29	1272.8	99.9	27755	119	US-60-495-114-16827	Sequence 16827, A
30	1272.8	99.9	37093	65	US-10-918-754-16929	Sequence 16929, A
31	1272.8	99.9	37093	119	US-60-495-114-16929	Sequence 16929, A
32	1272.8	99.9	37093	120	US-60-505-218-7506	Sequence 7506, Ap
33	1272.8	99.9	38141	114	US-60-449-629-854	Sequence 854, App
34	1272.8	99.9	42955	116	US-60-465-241-51333	Sequence 51333, A
35	1272.8	99.9	140599	65	US-10-918-754-16526	Sequence 16526, A
36	1272.8	99.9	140599	119	US-60-495-114-16526	Sequence 16526, A
37	1272.4	99.9	2829	48	US-10-108-260A-2899	Sequence 269, App
38	1254.4	98.5	2347	97	US-60-278-258-17389	Sequence 17389, A
39	1090.8	85.6	3070	120	US-60-505-218-276	Sequence 276, App
40	1061.4	83.3	3477	86	US-60-172-360-20587	Sequence 20587, A
41	1042.8	81.9	2440	58	US-10-641-643-588	Sequence 588, App
42	958	75.2	1298	33	US-09-724-676-6973	Sequence 6973, Ap
43	958	75.2	1298	33	US-09-724-676A-6973	Sequence 6973, Ap
44	958	75.2	1649	33	US-09-724-676-6970	Sequence 6970, Ap
45	958	75.2	1649	33	US-09-724-676A-6970	Sequence 6970, Ap

ALIGNMENTS

QY 721 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCATTAAG 780  
DB 721 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCATTAAG 780  
QY 781 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCATTAAG 840  
DB 781 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCATTAAG 840  
QY 841 GGTGCTGTTGACCTGAGGAGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTTCT 900  
DB 841 GGTGCTGTTGACCTGAGGAGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTTCT 900  
QY 901 GAGAACAGGAAATACAGATGCTCAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 960  
DB 901 GAGAACAGGAAATACAGATGCTCAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 960  
QY 961 ACACGCAAGCAAGCTGAAATTTTGTCTACAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1020  
DB 961 ACACGCAAGCAAGCTGAAATTTTGTCTACAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1020  
QY 1021 TTGACTTGGGAGGAGCTCTCTGATCTATTTGATGAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1080  
DB 1021 TTGACTTGGGAGGAGCTCTCTGATCTATTTGATGAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1080  
QY 1081 AGCGACTGTCTAGTCTTTGACACGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1140  
DB 1081 AGCGACTGTCTAGTCTTTGACACGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1140  
QY 1141 GGCATCAATGTAATCTATTTTCCATGTTGTAATGGCAATCAAAACATTTTCTGCGCCAGTGT 1200  
DB 1141 GGCATCAATGTAATCTATTTTCCATGTTGTAATGGCAATCAAAACATTTTCTGCGCCAGTGT 1200  
QY 1201 TTAAACTTCAAGTTCACAGAAATAGGACCCATCTGCTGSCCAACCTTAAACTCTTTT 1260  
DB 1201 TTAAACTTCAAGTTCACAGAAATAGGACCCATCTGCTGSCCAACCTTAAACTCTTTT 1260  
QY 1261 CGTAGTGGAGC 1274  
DB 1261 CGTAGTGGAGC 1274

RESULT 2  
US-10-679-246-1  
; Sequence 1, Application US/10679246  
; GENERAL INFORMATION:  
; APPLICANT: Matsuzawa, Shu-ichi  
; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved  
; FILE REFERENCE: 66821-235  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: US/09/591,694  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1274  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (274) ... (1167)  
US-10-679-246-1

Query Match 100.0%; Score 1274; DB 58; Length 1274;  
Best Local Similarity 100.0%; Pred. No. 8.9e-235;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTCTTTTGTAGTGTATGTCATTTTCTTATTTTGTAGCATTTATTTCTATGTAGTCTAT 60  
DB 1 TTCTTTTGTAGTGTATGTCATTTTCTTATTTTGTAGCATTTATTTCTATGTAGTCTAT 60

QY 61 CCNAGACGATTAAGGAGTTCACATGTTTTCGCGAACATTTTGAAGAGAGAGCTTATC 120  
DB 61 CCNAGACGATTAAGGAGTTCACATGTTTTCGCGAACATTTTGAAGAGAGAGCTTATC 120  
QY 121 CAGTGTACAGATCTTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT 180  
DB 121 CAGTGTACAGATCTTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT 180  
QY 181 TTAATCCTATTTTCTTCT 240  
DB 181 TTAATCCTATTTTCTTCT 240  
QY 241 TATGGCATGTAACATTAATTAATAAGTAAAGTAAAGTAAATTTTGAATTTTCTCTGCT 300  
DB 241 TATGGCATGTAACATTAATTAATAAGTAAAGTAAAGTAAATTTTGAATTTTCTCTGCT 300  
QY 301 CTTTATGTAATTTTATTTTACAGAAATGAGCGCTCAGATGCTACAGCATTTACCTACCGGTACC 360  
DB 301 CTTTATGTAATTTTATTTTACAGAAATGAGCGCTCAGATGCTACAGCATTTACCTACCGGTACC 360  
QY 361 TCGAAGTGTCCACCATTCAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 361 TCGAAGTGTCCACCATTCAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 421 GACTTGGCGAGTCTTTTTCAGATGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 421 GACTTGGCGAGTCTTTTTCAGATGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 481 CAATGTACAGAGGCGCATCTGTTTGTAGCAACTGTCGCCCAAGCTCAGATGCTGCTGCTGCT 540  
DB 481 CAATGTACAGAGGCGCATCTGTTTGTAGCAACTGTCGCCCAAGCTCAGATGCTGCTGCTGCT 540  
QY 541 ACTTGGCGGCGCTTTTGGGATCCATTCGCAACTTGGCTATGAGAGAAAGTGGCTAAATCA 600  
DB 541 ACTTGGCGGCGCTTTTGGGATCCATTCGCAACTTGGCTATGAGAGAAAGTGGCTAAATCA 600  
QY 601 GTACTTTTCCCTGTAAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 601 GTACTTTTCCCTGTAAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY 661 AAAGCAGACCATGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 661 AAAGCAGACCATGAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
QY 721 TCCTGTAATGGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATGCATCAGCATTAAG 780  
DB 721 TCCTGTAATGGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATGCATCAGCATTAAG 780  
QY 781 TCCTGTAATGGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATGCATCAGCATTAAG 840  
DB 781 TCCTGTAATGGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATGCATCAGCATTAAG 840  
QY 841 GGTGCTGTTGACCTGAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTTCTT 900  
DB 841 GGTGCTGTTGACCTGAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTTCTT 900  
QY 901 GAGAACAGGAAATACAGATGCTCAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 960  
DB 901 GAGAACAGGAAATACAGATGCTCAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 960  
QY 961 ACACGCAAGCAAGCTGAAATTTTGTCTACAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1020  
DB 961 ACACGCAAGCAAGCTGAAATTTTGTCTACAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1020  
QY 1021 TTGACTTGGGAGGAGCTCTCTGATCTATTTGATGAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1080  
DB 1021 TTGACTTGGGAGGAGCTCTCTGATCTATTTGATGAGGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1080  
QY 1081 AGCGACTGTCTAGTCTTTGACACGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1140  
DB 1081 AGCGACTGTCTAGTCTTTGACACGAGGATATAGTATTTTCTTGTCTACAGACATTAATCTT 1140  
QY 1141 GGCATCAATGTAATCTATTTTCCATGTTGTAATGGCAATCAAAACATTTTCTGCGCCAGTGT 1200

```
Db 1141 GGCATCAATGTAACATATTTCCATGTTGGAATGGCAATCAACATTTTCGGCCAGTGT 1200
QY 1201 TTAACATCTCAGTTTCACAGAAATAAGGCAACCATCTGTCTGGCAACCTAAACCTCTTT 1260
Db 1201 TTAACATCTCAGTTTCACAGAAATAAGGCAACCATCTGTCTGGCAACCTAAACCTCTTT 1260
QY 1261 CGGTAGTGGAGC 1274
Db 1261 CGGTAGTGGAGC 1274

RESULT 3
US-09-359-922-5961
; Sequence 5961, Application US/09359922
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/359,922
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: US 09/205,155
; EARLIER FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5961
; LENGTH: 1435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-359-922-5961

Query Match 100.0%; Score 1274; DB 21; Length 1435;
Best Local Similarity 100.0%; Pred. No. 9.1e-235;
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCTTTAGTTGTTATGGTCCATTTTCTATTATTAGCATTTTATTCTATGTAGTCTAT 60
Db 3 TTTCTTTAGTTGTTATGGTCCATTTTCTATTATTAGCATTTTATTCTATGTAGTCTAT 62
QY 61 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120
Db 63 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 122
QY 121 CAGGTACAGATCCTTAATAAGTGCAATTCAGTGTAAATTTTATTATTTTAAATATCTTT 180
Db 123 CAGGTACAGATCCTTAATAAGTGCAATTCAGTGTAAATTTTATTATTTTAAATATCTTT 182
QY 181 TTAATCCTATTTTCTCCTTTTCTGCTCAGTAAATTTTGTATGAAACTTTAAAGGACT 240
Db 183 TTAATCCTATTTTCTCCTTTTCTGCTCAGTAAATTTTGTATGAAACTTTAAAGGACT 242
QY 241 TATGGCATGTAACATATTATTAAGTAGTCAATGTTATATTAATTTTCTCCTGCCT 300
Db 243 TATGGCATGTAACATATTATTAAGTAGTCAATGTTATATTAATTTTCTCCTGCCT 302
QY 301 CCTTATGTTATTTTTCAGAAATGAGCCGTGACACTGCTACAGCAATTAACCTACCGGTACC 360
Db 303 CCTTATGTTATTTTTCAGAAATGAGCCGTGACACTGCTACAGCAATTAACCTACCGGTACC 362
QY 361 TCAGAGTGTCACATCCCAAGAGGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 363 TCAGAGTGTCACATCCCAAGAGGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
QY 421 GACTTCGCGAGTCTTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 423 GACTTCGCGAGTCTTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
QY 481 CAATGTACAGAGTGCCATCTTTTGTGTAGCAACTGTGCGCCAAAGCTCAATGTGTCCA 540
Db 483 CAATGTACAGAGTGCCATCTTTTGTGTAGCAACTGTGCGCCAAAGCTCAATGTGTCCA 542
QY 541 ACTTGGCGGGGCCCTTTGGGATCCAATTCGCACTTGGCTATGGAGAAAGTGCGCTAATCA 600
```

```
Db 543 ACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGTATGGAGAAAGTGCGCTAATTTCA 602
QY 601 GTACTTTTCCCTCTGAATATGCGTCTTCTGGATGTAATAAATCTCTGCCACACACAGAA 660
Db 603 GTACTTTTCCCTCTGAATATGCGTCTTCTGGATGTAATAAATCTCTGCCACACACAGAA 662
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTATAGGCTTTATTCCTGTCGCTGCCCTGGTCT 720
Db 663 AAAGCAGACCATGAAGAGCTCTGTGAGTTTATAGGCTTTATTCCTGTCGCTGCCCTGGTCT 722
QY 721 TCCTGTAATAAGGAGGCTCTGTGATGCTGTATATGCCCCATCTGATGCATCAGCATTAAG 780
Db 723 TCCTGTAATAAGGAGGCTCTGTGATGCTGTATATGCCCCATCTGATGCATCAGCATTAAG 782
QY 781 TCCATTACAACCTACAGGAGAGATATAGTTTTCTGTACAGACATTAATCTTCT 840
Db 783 TCCATTACAACCTACAGGAGAGATATAGTTTTCTGTACAGACATTAATCTTCT 842
QY 841 GGTGCTGTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTTCACTTCATGTTAGTCTTA 900
Db 843 GGTGCTGTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTTCACTTCATGTTAGTCTTA 902
QY 901 GAGAAACAGGAAATACAGTGGTCAACAGAGTCTTTCGCAATCGTACAGCTGATAGA 960
Db 903 GAGAAACAGGAAATACAGTGGTCAACAGAGTCTTTCGCAATCGTACAGCTGATAGA 962
QY 961 ACACGCAAGCAAGCTGAAAAATTTTCTTACCGACTTGAGCTAAATGGTCTATAGGCGACA 1020
Db 963 ACACGCAAGCAAGCTGAAAAATTTTCTTACCGACTTGAGCTAAATGGTCTATAGGCGACA 1022
QY 1021 TTGACTTGGGAAGCGACTCCTCGATCTATTCATGAAGGAATTTGCAAGCAATTAATGAAT 1080
Db 1023 TTGACTTGGGAAGCGACTCCTCGATCTATTCATGAAGGAATTTGCAAGCAATTAATGAAT 1082
QY 1081 AGCGACTGTCTAGTCTTTGACACACAGCATTTGACAGCTTTTTCGAGAAATGGCAATTTA 1140
Db 1083 AGCGACTGTCTAGTCTTTGACACACAGCATTTGACAGCTTTTTCGAGAAATGGCAATTTA 1142
QY 1141 GGCATCAATGTAACATATTTCATGTTGTAATGGCAATCAACATTTTTCGCCAGTGT 1200
Db 1143 GGCATCAATGTAACATATTTCATGTTGTAATGGCAATCAACATTTTTCGCCAGTGT 1202
QY 1201 TTAACATCTCAGTTTCACAGAAATAAGGCAACCATCTGTCTGGCAACCTAAACCTCTTT 1260
Db 1203 TTAACATCTCAGTTTCACAGAAATAAGGCAACCATCTGTCTGGCAACCTAAACCTCTTT 1262
QY 1261 CGGTAGTGGAGC 1274
Db 1263 CGGTAGTGGAGC 1276
```

```
RESULT 4
US-09-359-922-5961
; Sequence 5961, Application US/09359922A
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; TITLE OF INVENTION: LIBRARIES
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/359,922A
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: US 09/205,155
; EARLIER FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: US 09/034,341
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5961
; LENGTH: 1435
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-359-922-5961

Query Match 100.0%; Score 1274; DB 21; Length 1435;  
Best Local Similarity 100.0%; Pred. No. 9,1e-235;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCTTTAGTTGTTATGGTCCATTTCTTATTTTACCAATTTATTTCTATGTAGTCTAT 60  
DB 3 TTTCTTTAGTTGTTATGGTCCATTTCTTATTTTACCAATTTATTTCTATGTAGTCTAT 62

QY 61 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 120  
DB 63 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 122

QY 121 CAGGTACAGATCCTAATAAAGTGCAATTCAGTGTAATTTTATTTTAAATATCTTTT 180  
DB 123 CAGGTACAGATCCTAATAAAGTGCAATTCAGTGTAATTTTATTTTAAATATCTTTT 182

QY 181 TTAATCCTATTTTCTTCTCTCTTCTTCTCAGTAAATTTTGTATGAACCTTAAAGAGCT 240  
DB 183 TTAATCCTATTTTCTTCTCTCTTCTTCTCAGTAAATTTTGTATGAACCTTAAAGAGCT 242

QY 241 TATGGCATGTAAACATTTATTTAAGTAAGTCAAGTGTATTAATTTTCTTCTGCTCT 300  
DB 243 TATGGCATGTAAACATTTATTTAAGTAAGTCAAGTGTATTAATTTTCTTCTGCTCT 302

QY 301 CCTTATGTATTTATTTTCAAGAAATGAGCGGTGCTGCTACAGCAATTTACCTACCGGTACC 360  
DB 303 CCTTATGTATTTATTTTCAAGAAATGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCT 362

QY 361 TCGAAGTGTCCACCATCCACAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 363 TCGAAGTGTCCACCATCCACAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422

QY 421 GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTGTGCTGCTGCTGCTGCTGCTGCT 480  
DB 423 GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTGTGCTGCTGCTGCTGCTGCTGCT 482

QY 481 CAATGTACAGTGGCAATCTTTTGTAGTCAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
DB 483 CAATGTACAGTGGCAATCTTTTGTAGTCAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542

QY 541 ACTTGGCGGGGCTTTGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAATTTCA 600  
DB 543 ACTTGGCGGGGCTTTGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAATTTCA 602

QY 601 GTACTTTTCCCTGTAAATATGCTCTCTGATGTGAAATAACTCTGCTGCTGCTGCTGCTGCT 660  
DB 603 GTACTTTTCCCTGTAAATATGCTCTCTGATGTGAAATAACTCTGCTGCTGCTGCTGCTGCT 662

QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGGGCTTATTTCTGCTGCTGCTGCTGCTGCT 720  
DB 663 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGGGCTTATTTCTGCTGCTGCTGCTGCTGCT 722

QY 721 TCGTGTAAATGGGAAGGCTCTGTGATGTCTGTAATGCCCACTGTGATGATGATGATGATGAT 780  
DB 723 TCGTGTAAATGGGAAGGCTCTGTGATGTCTGTAATGCCCACTGTGATGATGATGATGATGAT 782

QY 781 TCCATTACAACTCAGGAGGAGGATATAGTTTCTTGTGCTACAGCAATTAATCTTCTCT 840  
DB 783 TCCATTACAACTCAGGAGGAGGATATAGTTTCTTGTGCTACAGCAATTAATCTTCTCT 842

QY 841 GGTGCTGTTGACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 843 GGTGCTGTTGACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902

QY 901 GAGAAACAGGAAATAACAGTGTCAACAGAGTCTTTCGCAATCGTACAGTGTATGAGA 960  
DB 903 GAGAAACAGGAAATAACAGTGTCAACAGAGTCTTTCGCAATCGTACAGTGTATGAGA 962

QY 961 ACACCGACAGCTGAAATTTTGTGCTTACCGACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 963 ACACCGACAGCTGAAATTTTGTGCTTACCGACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022

## RESULT 5

US-09-919-002-5961  
; Sequence S961, Application US/09919002  
; GENERAL INFORMATION:  
; APPLICANT: Leshkowitz, Dena  
; APPLICANT: Liu, Jin  
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
; TITLE OF INVENTION: LIBRARIES  
; FILE REFERENCE: 20411-752CON1  
; CURRENT APPLICATION NUMBER: US/09/919,002  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922  
; PRIOR FILING DATE: FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341  
; PRIOR FILING DATE: FILING DATE: 1998-02-13  
; NUMBER OF SEQ ID NOS: 13203  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5961  
; LENGTH: 1435  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-919-002-5961

Query Match 100.0%; Score 1274; DB 39; Length 1435;  
Best Local Similarity 100.0%; Pred. No. 9,1e-235;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCTTTAGTTGTTATGGTCCATTTCTTATTTTACCAATTTATTTCTATGTAGTCTAT 60  
DB 3 TTTCTTTAGTTGTTATGGTCCATTTCTTATTTTACCAATTTATTTCTATGTAGTCTAT 62

QY 61 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 120  
DB 63 CCAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 122

QY 121 CAGGTACAGATCCTAATAAAGTGCAATTCAGTGTAATTTTATTTTAAATATCTTTT 180  
DB 123 CAGGTACAGATCCTAATAAAGTGCAATTCAGTGTAATTTTATTTTAAATATCTTTT 182

QY 181 TTAATCCTATTTTCTTCTCTCTTCTTCTCAGTAAATTTTGTATGAACCTTAAAGAGCT 240  
DB 183 TTAATCCTATTTTCTTCTCTCTTCTTCTCAGTAAATTTTGTATGAACCTTAAAGAGCT 242

QY 241 TATGGCATGTAAACATTTATTTAAGTAAGTCAAGTGTATTAATTTTCTTCTGCTCT 300  
DB 243 TATGGCATGTAAACATTTATTTAAGTAAGTCAAGTGTATTAATTTTCTTCTGCTCT 302

QY 301 CCTTATGTATTTATTTTCAAGAAATGAGCGGTGCTGCTACAGCAATTTACCTACCGGTACC 360  
DB 303 CCTTATGTATTTATTTTCAAGAAATGAGCGGTGCTGCTACAGCAATTTACCTACCGGTACC 362



361 TCAGAGGTGTCACCAATCCAGAGGGTGCCTGCCCTGACTGCGCAACATGCAATCAACAAT 420  
363 TCAGAGGTGTCACCAATCCAGAGGGTGCCTGCCCTGACTGCGCAACATGCAATCAACAAT 422  
421 GACTGGCGAGTCTTTTGGAGTCCAGTCTGCTTTGACTATGTTTACCGCCCATCTTT 480  
423 GACTGGCGAGTCTTTTGGAGTCCAGTCTGCTTTGACTATGTTTACCGCCCATCTTT 482  
481 CAATGTCCAGAGTGGCATCTGTTTGTAGCAACTCTGCCCAAGCTCCACATGTTGTCCA 540  
483 CAATGTCCAGAGTGGCATCTGTTTGTAGCAACTCTGCCCAAGCTCCACATGTTGTCCA 542  
541 ACTTGGCGGGGCTTTGGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAAATCA 600  
543 ACTTGGCGGGGCTTTGGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAAATCA 602  
601 GTACTTTTCCCTGTAAATATATGCTCTCTGGATGTGAATAAATCTCTGCCACACAGAA 660  
603 GTACTTTTCCCTGTAAATATATGCTCTCTGGATGTGAATAAATCTCTGCCACACAGAA 662  
661 AAAGCAGACCATGAAGAGCTCTGTGAGTTAGGCTTATCTGTCCGTCGCTGCTGCT 720  
663 AAAGCAGACCATGAAGAGCTCTGTGAGTTAGGCTTATCTGTCCGTCGCTGCTGCT 722  
721 TCTGTAAATGGCAAGCTCTCTGGATGTGAATAAATCTCTGCCACACAGAA 780  
723 TCTGTAAATGGCAAGCTCTCTGGATGTGAATAAATCTCTGCCACACAGAA 782  
781 TCCATTACAAACCTTACAGGAGAGGATATAGTTTCTTGTACAGACATTAATCTTCT 840  
783 TCCATTACAAACCTTACAGGAGAGGATATAGTTTCTTGTACAGACATTAATCTTCT 842  
841 GGTGCTGTTGACTGGGTGATGAGTCTGTTTGGCTTTTCTTGTACAGACATTAATCTT 900  
843 GGTGCTGTTGACTGGGTGATGAGTCTGTTTGGCTTTTCTTGTACAGACATTAATCTT 902  
901 GAGAAACAGAAATACAGATGTCACAGAGTCTTCTGCAATCTGACAGCTGATAGGA 960  
903 GAGAAACAGAAATACAGATGTCACAGAGTCTTCTGCAATCTGACAGCTGATAGGA 962  
961 ACAGCAAGCAAGCTGAAATTTTGTCTACCACTTGGCTTAAATGGTATAGGCGACGA 1020  
963 ACAGCAAGCAAGCTGAAATTTTGTCTACCACTTGGCTTAAATGGTATAGGCGACGA 1022  
1021 TTGACTTGGAGAGGCTCTCGATCTTATTCATGAGGAATTCGAACAGCAATTAATGAAT 1080  
1023 TTGACTTGGAGAGGCTCTCGATCTTATTCATGAGGAATTCGAACAGCAATTAATGAAT 1082  
1081 AGCGACTGTCTAGTCTTTTGACACAGCATTTGCACAGCTTTTTCAGAAATGGCAATTTA 1140  
1083 AGCGACTGTCTAGTCTTTTGACACAGCATTTGCACAGCTTTTTCAGAAATGGCAATTTA 1142  
1141 GGCATCAATGTAATTTTCAATGTTGAATGGCAATCAACATTTTCTGCGCAGTGT 1200  
1143 GGCATCAATGTAATTTTCAATGTTGAATGGCAATCAACATTTTCTGCGCAGTGT 1202  
1201 TTAATACTTCAAGTTTACAGAAATAAGGCAACCTCTGTCTGCCAACCTTAAATCTTTT 1260  
1203 TTAATACTTCAAGTTTACAGAAATAAGGCAACCTCTGTCTGCCAACCTTAAATCTTTT 1262  
1261 CGGTAGGTGGAAGC 1274  
1263 CGGTAGGTGGAAGC 1276

RESULT 6  
US-09-644-867-5326  
; Sequence 5326, Application US/09644867  
; GENERAL INFORMATION:  
; APPLICANT: Holtzman, Douglas A.  
; APPLICANT: Donovan, Michael J.  
; APPLICANT: Leiby, Kevin R.  
; APPLICANT: Culpepper, Janice A.

Query Match 100.0%; Score 1274; DB 29; Length 1735;  
Best Local Similarity 100.0%; Pred. No. 9.4e-235;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
361 1 TTTCTTTAGTTGTTTATGGTCCATTTTCTATTTTAGCAATTTATTATTCTATGTAGTCTAT 60  
304 TTTCTTTAGTTGTTTATGGTCCATTTTCTATTTTAGCAATTTATTATTCTATGTAGTCTAT 363  
61 CCAAGACGATTAAGGAGATTCCACATGTTTTCGGGAACATTTTGAAGAAGAGCTTATC 120  
364 CCAAGACGATTAAGGAGATTCCACATGTTTTCGGGAACATTTTGAAGAAGAGCTTATC 423  
121 CAGTGTACAGATCCCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTT 180  
424 CAGTGTACAGATCCCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTT 483  
181 TTAATCCCTATTTTCTTCTCTCTTGTCTCAGTAAATTTTGTATGAAACTTTTAAAGGACT 240  
484 TTAATCCCTATTTTCTTCTCTCTTGTCTCAGTAAATTTTGTATGAAACTTTTAAAGGACT 543  
241 TATGTCATGTAAACATTTATTAAGTAAGTCAATGTTTATTTTCTCTGCTCT 300  
544 TATGTCATGTAAACATTTATTAAGTAAGTCAATGTTTATTTTCTCTGCTCT 603  
301 CTTTATGATTTTATTTTCAGAAATGAGCGTCACTGCTTACAGCAATTTACCTACCGGTACC 360  
604 CTTTATGATTTTATTTTCAGAAATGAGCGTCACTGCTTACAGCAATTTACCTACCGGTACC 663  
361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCCCTGACTGGCAAACTGCAATCCAAACAT 420  
664 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCCCTGACTGGCAAACTGCAATCCAAACAT 723  
421 GACTGGCGAGTCTTTTGGAGTCCAGTCTGCTTTGACTATGTTTACCGCCCATCTTT 480  
724 GACTGGCGAGTCTTTTGGAGTCCAGTCTGCTTTGACTATGTTTACCGCCCATCTTT 783  
481 CAAATGTCCAGTGGCCATCTTGTGTTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCA 540  
784 CAAATGTCCAGTGGCCATCTTGTGTTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCA 843  
541 ACTTGGCGGGGCTTTGGGATCCATTTGGGATCCATTTGGGATCCATTTGGGATCCATTT 903  
844 ACTTGGCGGGGCTTTGGGATCCATTTGGGATCCATTTGGGATCCATTTGGGATCCATTT 960  
601 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATAAATCTCTGCCACACAGAA 963  
904 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATAAATCTCTGCCACACAGAA 963  
661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCTTATTTCTGTCCGTCGCTGCTGCT 720  
964 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCTTATTTCTGTCCGTCGCTGCTGCT 1023  
721 TCTGTAAATGGCAAGGCTCTCTGGATGTGAATAAATCTCTGCCAACCTTAAATCTTTT 780

; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 1600.1173-001  
; CURRENT APPLICATION NUMBER: US/09/644,867  
; CURRENT FILING DATE: 2000-08-28  
; PRIOR APPLICATION NUMBER: 60/151,061  
; PRIOR FILING DATE: 1999-08-27  
; NUMBER OF SEQ ID NOS: 8090  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5326  
; LENGTH: 1735  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(1735)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-644-867-5326

Db 1024 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCATAG 1083  
Qy 781 TCATTACAACTCCTACAGGAGAGGATATAGTTTTCTTGCTACAGACATTAATCTTCCT 840  
Db 1084 TCATTACAACTCCTACAGGAGAGGATATAGTTTTCTTGCTACAGACATTAATCTTCCT 1143  
Qy 841 GGTGCTGTTGACTGGGCTGATGACAGTCTCTGTTTTGGCTTTCACATTCATGTAGTCTTA 900  
Db 1144 GGTGCTGTTGACTGGGCTGATGACAGTCTCTGTTTTGGCTTTCACATTCATGTAGTCTTA 1203  
Qy 901 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTCTTGCAATCGTACAGCTGATAGGA 960  
Db 1204 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTCTTGCAATCGTACAGCTGATAGGA 1263  
Qy 961 ACACGACAGAGCTGAAATTTTGTCTTACCCGACTTGAGCTAAATGGTCAATGGCGACGA 1020  
Db 1264 ACACGACAGAGCTGAAATTTTGTCTTACCCGACTTGAGCTAAATGGTCAATGGCGACGA 1323  
Qy 1021 TTGACTTGGAGAGGACTCTCGATCTATTTCATGAAGGAATGCAACAGCCATTATGAAT 1080  
Db 1324 TTGACTTGGAGAGGACTCTCGATCTATTTCATGAAGGAATGCAACAGCCATTATGAAT 1383  
Qy 1081 AGCGACTGTCTAGTCTTGACACACGACATTCGACAGCTTTTGCAGAAATGGCAATTTA 1140  
Db 1384 AGCGACTGTCTAGTCTTGACACACGACATTCGACAGCTTTTGCAGAAATGGCAATTTA 1443  
Qy 1141 GGCATCAATGTAATTTTCCATGTTTGAATGGCAATCAAACTTTTGGCCAGTGT 1200  
Db 1444 GGCATCAATGTAATTTTCCATGTTTGAATGGCAATCAAACTTTTGGCCAGTGT 1503  
Qy 1201 TTAATACCTTCACTTTCACAGAAATAGGACCCATCTGTCTGCCAACTTAAACTCTTT 1260  
Db 1504 TTAATACCTTCACTTTCACAGAAATAGGACCCATCTGTCTGCCAACTTAAACTCTTT 1563  
Qy 1261 CGGTAGTGGAGC 1274  
Db 1564 CGGTAGTGGAGC 1577

## RESULT 7 918-7745

US-09-652-918-7745  
; Sequence 7745, Application US/09652918  
; GENERAL INFORMATION:  
; APPLICANT: Galvin, Katherine  
; APPLICANT: Holtzman, Douglas A.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USRS  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 1600.1187-001  
; CURRENT APPLICATION NUMBER: US/09/652,918  
; CURRENT FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/151,130  
; PRIOR FILING DATE: 1999-08-30  
; NUMBER OF SEQ ID NOS: 8985  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7745  
; LENGTH: 1735  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(1735)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-652-918-7745

Query Match 100.0%; Score 1274; DB 29; Length 1735;  
Best Local Similarity 100.0%; Pred. No. 9,4e-235;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TTTCTTTAGTTGTTATGGTCCATTTTCTATTATTAGCATTTATTTCTATGTAGTCTAT 60  
Db 304 TTTCTTTAGTTGTTATGGTCCATTTTCTATTATTAGCATTTATTTCTATGTAGTCTAT 363  
Qy 61 CCAAGACGATTAAGGAGTTCACATGTTTTCGGAAACATTTTGAAGAGAGGCTTATC 120

Db 364 CCAAGACGATTAAGGAGTTCACATGTTTTCGGAAACATTTTGAAGAGAGGCTTATC 423  
Qy 121 CAGTGTACAGATCCTAATAAGTGCACATTCAGTGTAAATTTTATTTTAAATATCTTTT 180  
Db 424 CAGTGTACAGATCCTAATAAGTGCACATTCAGTGTAAATTTTATTTTAAATATCTTTT 483  
Qy 181 TTAATCCTATTTTCT 240  
Db 484 TTAATCCTATTTTCT 543  
Qy 241 TATGCGATGTAAACATTAATTTATAAGTAAGTGCATCTGTTATTAATTTTCTCTCTCT 300  
Db 544 TATGCGATGTAAACATTAATTTATAAGTAAGTGCATCTGTTATTAATTTTCTCTCTCT 603  
Qy 301 CTTTATGTATTTTATTTTTCAGAAATGAGCGCTCAGAGCTGTACAGCATTAACCTACCGGTACC 360  
Db 604 CTTTATGTATTTTATTTTTCAGAAATGAGCGCTCAGAGCTGTACAGCATTAACCTACCGGTACC 663  
Qy 361 TCGAAGTGTCCACATCCAGAGGGTGCCTGCGCTGACCTGGGCAAACTGCATCCAACT 420  
Db 664 TCGAAGTGTCCACATCCAGAGGGTGCCTGCGCTGACCTGGGCAAACTGCATCCAACT 723  
Qy 421 GACTTGGCGAGTCTTTTTCAGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCATCTT 480  
Db 724 GACTTGGCGAGTCTTTTTCAGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCATCTT 783  
Qy 481 CAATGTACAGTGGCCATCTTGTGTTAGCAACTGTGCGCCAAAGCTCAGATGTTGTCCA 540  
Db 784 CAATGTACAGTGGCCATCTTGTGTTAGCAACTGTGCGCCAAAGCTCAGATGTTGTCCA 843  
Qy 541 ACTTGGCGGGGCTTTTGGGATCCATTCGCAACTGTTGGCTATGGAGAAAGTGGCTAATCCA 600  
Db 844 ACTTGGCGGGGCTTTTGGGATCCATTCGCAACTGTTGGCTATGGAGAAAGTGGCTAATCCA 903  
Qy 601 GTACTTTCCCTGTAAATATGCGTCTTCTGATGTGAATTAACCTGCGCACACAGAA 660  
Db 904 GTACTTTCCCTGTAAATATGCGTCTTCTGATGTGAATTAACCTGCGCACACAGAA 963  
Qy 661 AAAGCAGACCATGAAGAGCTCTGTGATGTTTAGCGCTTATTCCTGTCGTCGCTGCTGTCT 720  
Db 964 AAAGCAGACCATGAAGAGCTCTGTGATGTTTAGCGCTTATTCCTGTCGTCGCTGCTGTCT 1023  
Qy 721 TCCTGTAATGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATGCATCAGCATAG 780  
Db 1024 TCCTGTAATGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATGCATCAGCATAG 1083  
Qy 781 TCCATTACACCCCTACAGGAGAGGATATAGTTTTTCTTGCTACAGACATTAATCTTCT 840  
Db 1084 TCCATTACACCCCTACAGGAGAGGATATAGTTTTTCTTGCTACAGACATTAATCTTCTCT 1143  
Qy 841 GGTGCTGTTGACTGGGTGATGATGATGCTCTGTTTTGGCTTTTCACTTCAATGTTAGTCTTA 900  
Db 1144 GGTGCTGTTGACTGGGTGATGATGATGCTCTGTTTTGGCTTTTCACTTCAATGTTAGTCTTA 1203  
Qy 901 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTTCCGCAATCGTACAGCTGATAGGA 960  
Db 1204 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTTCCGCAATCGTACAGCTGATAGGA 1263  
Qy 961 ACACGACAGAGCTGAAATTTTGTCTTACCCGACTTGAGCTAAATGGTCAATGGCGACGA 1020  
Db 1264 ACACGACAGAGCTGAAATTTTGTCTTACCCGACTTGAGCTAAATGGTCAATGGCGACGA 1323  
Qy 1021 TTGACTTGGAGAGGACTCTCTCGATCTATTTCATGAAGGAATGCAACAGCCATTATGAAT 1080  
Db 1324 TTGACTTGGAGAGGACTCTCTCGATCTATTTCATGAAGGAATGCAACAGCCATTATGAAT 1383  
Qy 1081 AGCGACTGTCTAGTCTTGACACACGACATTCGACAGCTTTTGCAGAAATGGCAATTTA 1140  
Db 1384 AGCGACTGTCTAGTCTTGACACACGACATTCGACAGCTTTTGCAGAAATGGCAATTTA 1443  
Qy 1141 GGCATCAATGTAATTTTCCATGTTTGAATGGCAATCAAACTTTTGGCCAGTGT 1200

Db 1444 GCATCAATGTAATATATTTCCATGTTGAAATGGCAATCAAAATTTTCTGGCCAGTGT 1503

Qy 1201 TTAACACTTCAGTTTCACAGAAATTAAGGACCCATCTGTCTGCGCAACCTTAAACTCTTTT 1260

Db 1504 TTAACACTTCAGTTTCACAGAAATTAAGGACCCATCTGTCTGCGCAACCTTAAACTCTTTT 1563

Qy 1261 CGTAGGTGGAAGC 1274

Db 1564 CGTAGGTGGAAGC 1577

RESULT 8

US-10-170-235-39760

; Sequence 39760, Application US/10170235

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig

; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN

; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF

; FILE REFERENCE: CL001380

; CURRENT APPLICATION NUMBER: US/10/170,235

; CURRENT FILING DATE: 2003-03-17

; NUMBER OF SEQ ID NOS: 42514

; SEQ ID NO 39760

; LENGTH: 2270

; TYPE: DNA

; ORGANISM: HUMAN

US-10-170-235-39760

Query Match 100.0%; Score 1274; DB 49; Length 2270;

Best Local Similarity 100.0%; Pred. No. 9.8e-235;

Mismatches 0; Conservative 0; Indels 0; Gaps 0;

Matches 1274;

Qy 1 TTTCTTTAGTTGTTATGTCCTCAATTTTCTATTTTAGCATTTATTTCTATGATGCTAT 60

Db 3 TTTCTTTAGTTGTTATGTCCTCAATTTTCTATTTTAGCATTTATTTCTATGATGCTAT 62

Qy 61 CCAAGACCAATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 120

Db 63 CCAAGACCAATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 122

Qy 121 CAGTGACAGATCCTAATAAAGTGCACTTCAGTGTAATTTTATTTTATTAATCTTTT 180

Db 123 CAGTGACAGATCCTAATAAAGTGCACTTCAGTGTAATTTTATTTTATTAATCTTTT 182

Qy 181 TTAATCCTATTTTCTTCT 240

Db 183 TTAATCCTATTTTCTTCT 242

Qy 241 TATGGCATGTAAACATTTATTTATAAGTAAGTCAATGTTTATTAATTTTCTCTGCTT 300

Db 243 TATGGCATGTAAACATTTATTTATAAGTAAGTCAATGTTTATTAATTTTCTCTGCTT 302

Qy 301 CTTATGTATTTATTTTCAAGAAATGAGCGGTGACATGCTACAGCATTTACCTACCGTACC 360

Db 303 CTTATGTATTTATTTTCAAGAAATGAGCGGTGACATGCTACAGCATTTACCTACCGTACC 362

Qy 361 TCGAGTGTCACCATCCAGAGGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

Db 363 TCGAGTGTCACCATCCAGAGGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422

Qy 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGTTTGTAGTGTCTGTTTGTAGTGTCTGTTT 480

Db 423 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGTTTGTAGTGTCTGTTTGTAGTGTCTGTTT 482

Qy 481 CAATGTCCAGATGCGCATCTTTTGTAGTGTCCAGTCTGTTTGTAGTGTCTGTTTGTAGTGTCT 540

Db 483 CAATGTCCAGATGCGCATCTTTTGTAGTGTCCAGTCTGTTTGTAGTGTCTGTTTGTAGTGTCT 542

Qy 541 ACTTGGCGGGGCTTTTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTC 600

Db 543 ACTTGGCGGGGCTTTTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTC 602

Qy 601 GTACTTTTCCCTGTAAATATGTCGTCTTCTGGATGTGAAATTAATCTGTGCCACACAGAA 660

Db 603 GTACTTTTCCCTGTAAATATGCGCTCTTCTGGATGTGAAATTAATCTGTGCCACACAGAA 662

Qy 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGAAGGCTTATTTCTGTCCGTGCGCTGGTGTCT 720

Db 663 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGAAGGCTTATTTCTGTCCGTGCGCTGGTGTCT 722

Qy 721 TCCTGTAATGGCAAGGCTCTCTGGATGTCTTAATGCCCCCATCTGTGATGCATCAGCATAAAG 780

Db 723 TCCTGTAATGGCAAGGCTCTCTGGATGTCTTAATGCCCCCATCTGTGATGCATCAGCATAAAG 782

Qy 781 TCCATTACAACCCCTACAGGAGAGATATAGTTTCTTGTCTGTACAGACATTAATCTTCTT 840

Db 783 TCCATTACAACCCCTACAGGAGAGATATAGTTTCTTGTCTGTACAGACATTAATCTTCTT 842

Qy 841 GGTGCTGTGATGGTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900

Db 843 GGTGCTGTGATGGTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902

Qy 901 GAGAAACAGGAAAAATACGATGTCACCCAGCAGTTCTTTCGCAATCGTACAGCTGATAGGA 960

Db 903 GAGAAACAGGAAAAATACGATGTCACCCAGCAGTTCTTTCGCAATCGTACAGCTGATAGGA 962

Qy 961 ACACGCAAGCAAGCTGAAAAATTTTCTTACCGACTTGAGCTTAATGCTCATAGGCGACGA 1020

Db 963 ACACGCAAGCAAGCTGAAAAATTTTCTTACCGACTTGAGCTTAATGCTCATAGGCGACGA 1022

Qy 1021 TTGACTTCGGAAGCGACTCCCTCGATCTATTATGAAGAAATTCGCAACAGCCATTTATGAAT 1080

Db 1023 TTGACTTCGGAAGCGACTCCCTCGATCTATTATGAAGAAATTCGCAACAGCCATTTATGAAT 1082

Qy 1081 AGCGACTCTAGTCTTTTGACACCCAGCATGTCAGCTTTTTCGCAAAATGCGCAATTTA 1140

Db 1083 AGCGACTCTAGTCTTTTGACACCCAGCATGTCAGCTTTTTCGCAAAATGCGCAATTTA 1142

Qy 1141 GCATCAATGTAATCTTTTCCATGTTTGAATGCAATCAAAATTTTCTGGCCAGTGT 1200

Db 1143 GCATCAATGTAATCTTTTCCATGTTTGAATGCAATCAAAATTTTCTGGCCAGTGT 1202

Qy 1201 TTAACACTTCAGTTTCACAGAAATTAAGGACCCATCTGTCTGCGCAACCTTAAACTCTTT 1260

Db 1203 TTAACACTTCAGTTTCACAGAAATTAAGGACCCATCTGTCTGCGCAACCTTAAACTCTTT 1262

Qy 1261 CGTAGGTGGAAGC 1274

Db 1263 CGTAGGTGGAAGC 1276

RESULT 9

US-60-324-185-25086

; Sequence 25086, Application US/60324185

; GENERAL INFORMATION:

; APPLICANT: Morris, MacDonal

; APPLICANT: Lal, Preeti

; APPLICANT: Diep, Binh

; TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING

; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE

; TITLE OF INVENTION: POLYMORPHISMS IDENTIFIED THEREBY

; FILE REFERENCE: GX-0019-1 P

; CURRENT APPLICATION NUMBER: US/60/324,185

; CURRENT FILING DATE: 2001-09-21

; NUMBER OF SEQ ID NOS: 35862

; SOFTWARE: PERL Program

; SEQ ID NO 25086

; LENGTH: 2599

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No: 348851.1

; NAME/KEY: unsure

; LOCATION: 2470

; OTHER INFORMATION: a, t, c, g, or other

US-60-324-185-25086

Query Match 100.0%; Score 1274; DB 102; Length 2599;  
Best Local Similarity 100.0%; Pred. No. 1e-234;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TTTCTTTAGTTGTTATGTTCCATTTTCTATTATTAGCATTTATTATTCTATGTTAGTCTAT 60
DB 198 TTTCTTTAGTTGTTATGTTCCATTTTCTATTATTAGCATTTATTATTCTATGTTAGTCTAT 257

QY 61 CCAAGACGATTAAAGGAGTTCACATGTTTTCGGAACAATTTTGAAGAGAGAGCTTATC 120
DB 258 CCAAGACGATTAAAGGAGTTCACATGTTTTCGGAACAATTTTGAAGAGAGAGCTTATC 317

QY 121 CAGTGTACAGATCCTTAATAAGTGCAATTCAGTGTAATTTTATTTTAAATATCTTTT 180
DB 318 CAGTGTACAGATCCTTAATAAGTGCAATTCAGTGTAATTTTATTTTAAATATCTTTT 377

QY 181 TTAATCCTATTTTCTCTCTCTCTCTCTCAGTAAATTTTGTATGAACCTTTAAAGGACT 240
DB 378 TTAATCCTATTTTCTCTCTCTCTCTCTCAGTAAATTTTGTATGAACCTTTAAAGGACT 437

QY 241 TATGCAATGTAACATTTATTATAAGTAAGTCATGTTTATATTAATTTTCTCTGCT 300
DB 438 TATGCAATGTAACATTTATTATAAGTAAGTCATGTTTATATTAATTTTCTCTGCT 497

QY 301 CTTTATGTTATTTTTCAGAAATGAGCCGTGACGTGTACAGCAATTTACCTACCGTACC 360
DB 498 CTTTATGTTATTTTTCAGAAATGAGCCGTGACGTGTACAGCAATTTACCTACCGTACC 557

QY 361 TCGAAGTGTCCACATCCACAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 558 TCGAAGTGTCCACATCCACAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617

QY 421 GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTTGTACCGCCATCTT 480
DB 618 GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTTGTACCGCCATCTT 677

QY 481 CAATGTACAGTGGCATCTTGTGTTAGCAACTGTCGCCAAAGCTCATGTTGTCCA 540
DB 678 CAATGTACAGTGGCATCTTGTGTTAGCAACTGTCGCCAAAGCTCATGTTGTCCA 737

QY 541 ACTTGGCGGGCCCTTTGGATGCCATTCGCACTTGGCTATGAGAAAGTGGCTAATCCA 600
DB 738 ACTTGGCGGGCCCTTTGGATGCCATTCGCACTTGGCTATGAGAAAGTGGCTAATCCA 797

QY 601 GTACTTTTCCCTGTAAATATGCTCTCTGATGTGAATAACTCTGCACACACAGAA 660
DB 798 GTACTTTTCCCTGTAAATATGCTCTCTGATGTGAATAACTCTGCACACACAGAA 857

QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCTTTATTCCTGTCGTCGCTGCTGCT 720
DB 858 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCTTTATTCCTGTCGTCGCTGCTGCT 917

QY 721 TCGTGTAAATGGAGGCTCTCTGATGCTGTAAATGCCCTCTGATGATGATGATGATGAT 780
DB 918 TCGTGTAAATGGAGGCTCTCTGATGCTGTAAATGCCCTCTGATGATGATGATGATGAT 977

QY 781 TCCATTACAACTCAGGAGGAGGATATAGTTTTCCTGCTACAGACATTAATCTTCT 840
DB 978 TCCATTACAACTCAGGAGGAGGATATAGTTTTCCTGCTACAGACATTAATCTTCT 1037

QY 841 GGTGCTGTTGACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1038 GGTGCTGTTGACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1097

QY 901 GAGAAACAGGAAATAACGATGGTCAACAGAGTCTTCGCAATCGTACAGCTGATGAGA 960
DB 1098 GAGAAACAGGAAATAACGATGGTCAACAGAGTCTTCGCAATCGTACAGCTGATGAGA 1157

QY 961 ACACGCAAGCACTGAAATTTTGTCTTACCGACTTGAATGATGATGATGATGATGATGAT 1020
DB 1158 ACACGCAAGCACTGAAATTTTGTCTTACCGACTTGAATGATGATGATGATGATGATGAT 1217
```

RESULT 10

US-09-644-867-7203

; Sequence 7203; Application US/09644867

; GENERAL INFORMATION:

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Donovan, Michael J.

; APPLICANT: Leiby, Kevin R.

; APPLICANT: Culpepper, Janice A.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

; FILE OF INVENTION: THEREFOR

; FILE REFERENCE: 1600.1173-001

; CURRENT APPLICATION NUMBER: US/09/644,867

; PRIOR FILING DATE: 2000-08-28

; PRIOR APPLICATION NUMBER: 60/151,061

; PRIOR FILING DATE: 1999-08-27

; NUMBER OF SEQ ID NOS: 8090

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7203

; LENGTH: 2921

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)..(2921)

; OTHER INFORMATION: n = A,T,C or G

US-09-644-867-7203

Query Match 100.0%; Score 1274; DB 29; Length 2921;

Best Local Similarity 100.0%; Pred. No. 1e-234;

Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TTTCTTTAGTTGTTATGTTCCATTTTCTATTATTAGCATTTATTATTCTATGTTAGTCTAT 60
DB 479 TTTCTTTAGTTGTTATGTTCCATTTTCTATTATTAGCATTTATTATTCTATGTTAGTCTAT 538

QY 61 CCAAGACGATTAAAGGAGTTCACATGTTTTCGGAACAATTTTGAAGAGAGAGCTTATC 120
DB 539 CCAAGACGATTAAAGGAGTTCACATGTTTTCGGAACAATTTTGAAGAGAGAGCTTATC 598

QY 121 CAGTGTACAGATCCTTAATAAGTGCAATTCAGTGTAAATTTTATTTTAAATATCTTTT 180
DB 599 CAGTGTACAGATCCTTAATAAGTGCAATTCAGTGTAAATTTTATTTTAAATATCTTTT 658

QY 181 TTAATCCTATTTTCTCTCTCTCTCTCAGTAAATTTTGTATGAACCTTTAAAGGACT 240
DB 659 TTAATCCTATTTTCTCTCTCTCTCTCAGTAAATTTTGTATGAACCTTTAAAGGACT 718

QY 241 TATGCAATGTAACATTTATTATAAGTAAGTCATGTTTATATTAATTTTCTCTGCT 300
DB 719 TATGCAATGTAACATTTATTATAAGTAAGTCATGTTTATATTAATTTTCTCTGCT 778
```



Db 1139 AAGCAGACCATGAAGAGCTCTGTAGTTAGGCCCTTATTCCTGTCGGTCCCTGTGTCT 1198  
Qy 721 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGATGCATCAGCATAAG 780  
Db 1199 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGATGCATCAGCATAAG 1258  
Qy 781 TCATTTACAACCTACAGGAGAGGATATAGTTTTCTTGTCTACAGACATTAATCTTCTCT 840  
Db 1259 TCATTTACAACCTACAGGAGAGGATATAGTTTTCTTGTCTACAGACATTAATCTTCTCT 1318  
Qy 841 GGTGCTGTTGACTGGGTGATGATGCAAGTCTGTTTGGCTTTCACATTTCACTTGTAGTCTTA 900  
Db 1319 GGTGCTGTTGACTGGGTGATGATGCAAGTCTGTTTGGCTTTCACATTTCACTTGTAGTCTTA 1378  
Qy 901 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 960  
Db 1379 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 1438  
Qy 961 ACAGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGTCATAGGCGACGA 1020  
Db 1439 ACAGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGTCATAGGCGACGA 1498  
Qy 1021 TTGACTTTGGAGCGACTCTCTCGATCTATTATGAAAGGAATTGCAACAGCCATTATGAAT 1080  
Db 1499 TTGACTTTGGAGCGACTCTCTCGATCTATTATGAAAGGAATTGCAACAGCCATTATGAAT 1558  
Qy 1081 AGCGACTGTCTAGTCTTTGACACAGCATTGCAACAGCTTTTTCAGAAAAATGGCAATTTA 1140  
Db 1559 AGCGACTGTCTAGTCTTTGACACAGCATTGCAACAGCTTTTTCAGAAAAATGGCAATTTA 1618  
Qy 1141 GGCATCAATGTAATTTTCCATGTTGGAATGGCAATCAACATTTTTCGCCCAGTGT 1200  
Db 1619 GGCATCAATGTAATTTTCCATGTTGGAATGGCAATCAACATTTTTCGCCCAGTGT 1678  
Qy 1201 TTAAATCTTCAGTTTCACAGAAATAGGCACCCATCTGCTGCCAACCTTAAATCTTTT 1260  
Db 1679 TTAAATCTTCAGTTTCACAGAAATAGGCACCCATCTGCTGCCAACCTTAAATCTTTT 1738  
Qy 1261 CGGTAGGTGGAGC 1274  
Db 1739 CGGTAGGTGGAGC 1752

## RESULT 12

US-09-726-811-5469  
; Sequence 5469, Application US/09726811  
; GENERAL INFORMATION:  
; APPLICANT: Gutierrez-Ramos, Jose-Carlos  
; APPLICANT: Wen, Danyi  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; FILE REFERENCE: 1600-2027-001  
; CURRENT APPLICATION NUMBER: US/09/726,811  
; CURRENT FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/168,136  
; PRIOR FILING DATE: 1999-11-30  
; NUMBER OF SEQ ID NOS: 5515  
; SOFTWARE: Pstseq for Windows Version 4.0  
; SEQ ID NO 5469  
; LENGTH: 2921  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(2921)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-726-811-5469

## Query Match

100.0%; Score 1274; DB 33; Length 2921;

Best Local Similarity 100.0%; Pred. No. 1e-234;

Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTCTTTAGTTGTTTATGCTCCATTTTCTTATTTTACGATTTATTTCTATGTAGTCTAT 60  
Db 479 TTTCTTTAGTTGTTTATGCTCCATTTTCTTATTTTACGATTTATTTCTATGTAGTCTAT 538  
Qy 61 CCAGAGACCATTAAGGAGTTCACATGTTTTCGGAACATTTTGAAGAGAGAGCTTATC 120  
Db 539 CCAGAGACCATTAAGGAGTTCACATGTTTTCGGAACATTTTGAAGAGAGAGCTTATC 598  
Qy 121 CAGTGTACAGATCCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATCTTTT 180  
Db 599 CAGTGTACAGATCCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATCTTTT 658  
Qy 181 TTAATCTTATTTTCTTCTCTTTTGTCTCAGTAAATTTTGTATGAAATCTTTAAAGAGCT 240  
Db 659 TTAATCTTATTTTCTTCTCTTTTGTCTCAGTAAATTTTGTATGAAATCTTTAAAGAGCT 718  
Qy 241 TATGGCATGTAACATTTATTAAGTAAGTATGCTGTTTATTAATTTTCTCTCCCT 300  
Db 719 TATGGCATGTAACATTTATTAAGTAAGTATGCTGTTTATTAATTTTCTCTCCCT 778  
Qy 301 CCTTATGTTATTTTTCAGAAATGAGCCCTCAGACTGCTACAGCATTTACCTACCGGTACC 360  
Db 779 CCTTATGTTATTTTTCAGAAATGAGCCCTCAGACTGCTACAGCATTTACCTACCGGTACC 838  
Qy 361 TCGAAGTGTCCACATCCAGAGGGTCTGCCCTGACTGGCAAACTGCAATCCAAAT 420  
Db 839 TCGAAGTGTCCACATCCAGAGGGTCTGCCCTGACTGGCAAACTGCAATCCAAAT 898  
Qy 421 GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTGTGACTGTGTTTACCGCCATCTT 480  
Db 899 GACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGCTTGTGACTGTGTTTACCGCCATCTT 958  
Qy 481 CAATGTACAGTGGGCTCTTCTTGTAGCAACTGTGCCCAAGCTCACATGTTCTCCA 540  
Db 959 CAATGTACAGTGGGCTCTTCTTGTAGCAACTGTGCCCAAGCTCACATGTTCTCCA 1018  
Qy 541 ACTTGGCGGGGCTCTTGGGATCCATTTGGCACTTGGCTATGGAGAAAGTGGCTAAATCA 600  
Db 1019 ACTTGGCGGGGCTCTTGGGATCCATTTGGCACTTGGCTATGGAGAAAGTGGCTAAATCA 1078  
Qy 601 GTACTTTTCCCTGTAAATATGCTTCTGATGTAATTAATCTTGCACACACAGAA 660  
Db 1079 GTACTTTTCCCTGTAAATATGCTTCTGATGTAATTAATCTTGCACACACAGAA 1138  
Qy 661 AAAGCAGACCATTAAGAGCTCTGTGAGTTTAGGCCCTTATTTCTGTCGGTCCCTGTGTCT 720  
Db 1139 AAAGCAGACCATTAAGAGCTCTGTGAGTTTAGGCCCTTATTTCTGTCGGTCCCTGTGTCT 1198  
Qy 721 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCATAAG 780  
Db 1199 TCCTGTAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCATAAG 1258  
Qy 781 TCATTTACACCTTACAGGAGAGGATATAGTTTTTCTTGTCTACAGACATTAATCTTCTCT 840  
Db 1259 TCATTTACACCTTACAGGAGAGGATATAGTTTTTCTTGTCTACAGACATTAATCTTCTCT 1318  
Qy 841 GGTGCTGTTGACTGGGTGATGATGCAAGTCTGCTGTTTGGCTTTCACATTTCACTTGTAGTCTTA 900  
Db 1319 GGTGCTGTTGACTGGGTGATGATGCAAGTCTGCTGTTTGGCTTTCACATTTCACTTGTAGTCTTA 1378  
Qy 901 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 960  
Db 1379 GAGAAACAGGAAATACGATGCTCACCAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 1438  
Qy 961 ACAGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGTCATAGGCGACGA 1020  
Db 1439 ACAGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGTCATAGGCGACGA 1498  
Qy 1021 TTGACTTTGGAGCGACTCTCTCGATCTATTATGAAAGGAATTGCAACAGCCATTATGAAT 1080  
Db 1499 TTGACTTTGGAGCGACTCTCTCGATCTATTATGAAAGGAATTGCAACAGCCATTATGAAT 1558  
Qy 1081 AGCGACTGTCTAGTCTTTGACACAGCATTGCAACAGCTTTTTCAGAAAAATGGCAATTTA 1140







```
RESULT 14
US-09-785-276A-25457
; Sequence 25457, Application US/09785276A
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007B
; CURRENT APPLICATION NUMBER: US/09/785,276A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25457
; LENGTH: 2924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 57, 2923, 2924
; OTHER INFORMATION: n = A,T,C or G
US-09-785-276A-25457.

Query Match      100.0%; Score 1274; DB 34; Length 2924;
Best Local Similarity 100.0%; Pred. No. 1e-234;
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  TTCTTTAGCTGTTATGTCCTCAATTTTCTATTTAGCAATTTATTTCTATGCTAGTCTAT 60
DB      480 TTCTTTAGCTGTTATGTCCTCAATTTTCTATTTAGCAATTTATTTCTATGCTAGTCTAT 539

QY      61  CCAAGACGATTAAGGGAGTTCACATGTTTCCGGAAACATTTTGAAGAGAGAGCTTATC 120
DB      540 CCAAGACGATTAAGGGAGTTCACATGTTTCCGGAAACATTTTGAAGAGAGAGCTTATC 599

QY      121 CAGTGACAGATCCTAATAAGTGACATTCAGTGTAATTTTATTTTATTAATATCTTTT 180
DB      600 CAGTGACAGATCCTAATAAGTGACATTCAGTGTAATTTTATTTTATTAATATCTTTT 659

QY      181 TTAATCCTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
DB      660 TTAATCCTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 719

QY      241 TATGGCATGTAACATTTATTTATAAGTAAGTCATGTTATTAATTTTCTCTCTCTCT 300
DB      720 TATGGCATGTAACATTTATTTATAAGTAAGTCATGTTATTAATTTTCTCTCTCTCT 779

QY      301 CCTTATGTTATTTATTTTCAAGAAATGAGCCGTGACATGCTGACAGCAATTTACCTACCGGTACC 360
DB      780 CCTTATGTTATTTATTTTCAAGAAATGAGCCGTGACATGCTGACAGCAATTTACCTACCGGTACC 839

QY      361 TCGAAGTGTCACCATCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB      840 TCGAAGTGTCACCATCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899

QY      421 GACTTGGCCAGTCTTTTTCAGTGTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB      900 GACTTGGCCAGTCTTTTTCAGTGTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
```

```
QY      481 CAATGTCAGATGGCCATCTTGTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCA 540
DB      960 CAATGTCAGATGGCCATCTTGTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCA 1019

QY      541 ACTTGGCCGGGCGCTTTGGGATCCATTTCGCAACTTGGCTATGAGAAAGTGGCTAATTTCA 600
DB      1020 ACTTGGCCGGGCGCTTTGGGATCCATTTCGCAACTTGGCTATGAGAAAGTGGCTAATTTCA 1079

QY      601 GTACTTTTCCCTGTAAATATATGCGTCTTCTGGATGTGAATAAATCTGTGCCACACACAGAA 660
DB      1080 GTACTTTTCCCTGTAAATATATGCGTCTTCTGGATGTGAATAAATCTGTGCCACACACAGAA 1139

QY      661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGCCCTTATTCCTGTCGCTGCTGCTGCT 720
DB      1140 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGCCCTTATTCCTGTCGCTGCTGCTGCT 1199

QY      721 TCTGTAAATGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGATGATCAGCATAAG 780
DB      1200 TCTGTAAATGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGATGATCAGCATAAG 1259

QY      781 TCCATTACAAACCTACAGGGAGAGATATAGTTTTTCTTGTCTACAGACATTAATCTTCCT 840
DB      1260 TCCATTACAAACCTACAGGGAGAGATATAGTTTTTCTTGTCTACAGACATTAATCTTCCT 1319

QY      841 GGTGCTGTTGACTGGGTGATGATGCTGCTGTTTGGCTTTTCACTTATGTTAGTCTTA 900
DB      1320 GGTGCTGTTGACTGGGTGATGATGCTGCTGTTTGGCTTTTCACTTATGTTAGTCTTA 1379

QY      901 GAGAAACAGGAAATAATACATGCTCACCAGCAGTCTTCCGCAATCGTACAGCTGATAGGA 960
DB      1380 GAGAAACAGGAAATAATACATGCTCACCAGCAGTCTTCCGCAATCGTACAGCTGATAGGA 1439

QY      961 ACACGCAAGCAAGCTGAAATTTTGTCTACCGACTTTGAGCTAAATGGTCAATAGGCGACGA 1020
DB      1440 ACACGCAAGCAAGCTGAAATTTTGTCTACCGACTTTGAGCTAAATGGTCAATAGGCGACGA 1499

QY      1021 TTGACTTGGGAAAGCGACTCCTCGATCTATTCATGAAGAAATTTGCAACAGCATTATGAAT 1080
DB      1500 TTGACTTGGGAAAGCGACTCCTCGATCTATTCATGAAGAAATTTGCAACAGCATTATGAAT 1559

QY      1081 AGCGACTGTCTAGTCTTTTGACACACAGCATTGACAGCTTTTTCGAGAAATGCGCAATTTTA 1140
DB      1560 AGCGACTGTCTAGTCTTTTGACACACAGCATTGACAGCTTTTTCGAGAAATGCGCAATTTTA 1619

QY      1141 GGCATCAATGTAACATTTTCCATGTTTGAATGGCAATCAAACTTTTCTGCGCCAGTGT 1200
DB      1620 GGCATCAATGTAACATTTTCCATGTTTGAATGGCAATCAAACTTTTCTGCGCCAGTGT 1679

QY      1201 TTAACACTTCAGTTTTCAGAGAAATAAGGACCCATCTGCTGCGCAACCTAAACCTCTTT 1260
DB      1680 TTAACACTTCAGTTTTCAGAGAAATAAGGACCCATCTGCTGCGCAACCTAAACCTCTTT 1739

QY      1261 CGGTAGGTGGAAGC 1274
DB      1740 CGGTAGGTGGAAGC 1753

RESULT 15
US-09-785-276A-28822
; Sequence 28822, Application US/09785276A
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007B
; CURRENT APPLICATION NUMBER: US/09/785,276A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
```

;; PRIOR FILING DATE: 2000-03-16  
;; PRIOR APPLICATION NUMBER: 60/207,454  
;; PRIOR FILING DATE: 2000-05-25  
;; PRIOR APPLICATION NUMBER: 60/211,314  
;; PRIOR FILING DATE: 2000-06-09  
;; PRIOR APPLICATION NUMBER: 60/219,007  
;; PRIOR FILING DATE: 2000-07-18  
;; PRIOR APPLICATION NUMBER: 60/255,281  
;; PRIOR FILING DATE: 2000-12-13  
;; NUMBER OF SEQ ID NOS: 62332  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 28822  
;; LENGTH: 2924  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: 1, 57, 2923, 2924  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-785-276A-28822

Query Match 100.0%; Score 1274; DB 34; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 1e-234;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTTAGTGTATGTCCTATTTTCTATTTAGCATTTTATTTCTATGTTAGTCTAT 60  
DB 480 TTTCTTTAGTGTATGTCCTATTTTCTATTTAGCATTTTATTTCTATGTTAGTCTAT 539  
QY 61 CCAAGACGATTAAGGAGTTCACATGTTTTCGGAACATTTTGAAGAGAGAGCTTATC 120  
DB 540 CCAAGACGATTAAGGAGTTCACATGTTTTCGGAACATTTTGAAGAGAGAGCTTATC 599  
QY 121 CAGGTACAGATCCTAATAAGTGACATTCAGTGAATTTTATTTTATATATCTTTT 180  
DB 600 CAGGTACAGATCCTAATAAGTGACATTCAGTGAATTTTATTTTATATATCTTTT 659  
QY 181 TTAATCCTATTTTCTTCTCTTCTGCTAGTAAATTTTGTATGAACTTTAAAGAGACT 240  
DB 660 TTAATCCTATTTTCTTCTCTTCTGCTAGTAAATTTTGTATGAACTTTAAAGAGACT 719  
QY 241 TATGGCATGTAACATTTATTAAGTAAGTCAATGTTATATATTTTCTCTGCTCT 300  
DB 720 TATGGCATGTAACATTTATTAAGTAAGTCAATGTTATATATTTTCTCTGCTCT 779  
QY 301 CCTATGATTTATTTTTCAGAAATGAGCGTCACTGCTACAGCATTTACCTCCGCTACC 360  
DB 780 CCTATGATTTATTTTTCAGAAATGAGCGTCACTGCTACAGCATTTACCTCCGCTACC 839  
QY 361 TCGAAGTGTCCACCATCCAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 840 TCGAAGTGTCCACCATCCAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899  
QY 421 GACTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 900 GACTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959  
QY 481 CAATGTCAGAGTGGCCATTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
DB 960 CAATGTCAGAGTGGCCATTTTGTAGTGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1019  
QY 541 ACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATTC 600  
DB 1020 ACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATTC 1079  
QY 601 GTACTTTTCCCTGTAAATATAGTGTCTTCTGGATGTGAATTAATCTGCCCACACAGAA 660  
DB 1080 GTACTTTTCCCTGTAAATATAGTGTCTTCTGGATGTGAATTAATCTGCCCACACAGAA 1139  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTAGGCTTATTTCTGTCGCTGCTGCTGCTGCT 720  
DB 1140 AAAGCAGACCATGAAGAGCTCTGTGAGTTAGGCTTATTTCTGTCGCTGCTGCTGCTGCT 1199

QY 721 TCCTGTAATGCAAGGCTCTCTGATGCTGTAAATGCCCATCTGATGATCATCATCAAG 780  
DB 1200 TCCTGTAATGCAAGGCTCTCTGATGCTGTAAATGCCCATCTGATGATCATCATCAAG 1259  
QY 781 TCCATTACAACCCCTACAGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCT 840  
DB 1260 TCCATTACAACCCCTACAGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCT 1319  
QY 841 GGTGCTGTGATCTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
DB 1320 GGTGCTGTGATCTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379  
QY 901 GAGAAACAGGAAATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
DB 1380 GAGAAACAGGAAATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1439  
QY 961 ACACGCAAGCAAGCTGAAATTTTCTTACCGACTTGAGCTTAAATGCTCATAGGCGAGA 1020  
DB 1440 ACACGCAAGCAAGCTGAAATTTTCTTACCGACTTGAGCTTAAATGCTCATAGGCGAGA 1499  
QY 1021 TTGACTTGGGAAGCGACTCTCTCGATCTATTTTCAATGAAAGAAATGCAACAGCCATTATGAAT 1080  
DB 1500 TTGACTTGGGAAGCGACTCTCTCGATCTATTTTCAATGAAAGAAATGCAACAGCCATTATGAAT 1559  
QY 1081 AGCGACTGTCTAGTCTTTTGACACAGCATTTGACAGCTTTTTCGAGAAATGCGCAATTTA 1140  
DB 1560 AGCGACTGTCTAGTCTTTTGACACAGCATTTGACAGCTTTTTCGAGAAATGCGCAATTTA 1619  
QY 1141 GGCATCAATGTAACATTTTCCATGTTGAAATGCAATCAAACTTTTCTGGCCAGTGT 1200  
DB 1620 GGCATCAATGTAACATTTTCCATGTTGAAATGCAATCAAACTTTTCTGGCCAGTGT 1679  
QY 1201 TTAAAACTTCAGTTTTCACAGAAATTAAGGCACTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
DB 1680 TTAAAACTTCAGTTTTCACAGAAATTAAGGCACTGCTGCTGCTGCTGCTGCTGCTGCT 1739  
QY 1261 CGGTAGGTGGAAGC 1274  
DB 1740 CGGTAGGTGGAAGC 1753

Search completed: April 25, 2005, 00:27:40  
Job time : 5412 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 20:02:16 ; Search time 474 Seconds  
(without alignments)  
4245.569 Million cell updates/sec

Title: US-10-679-246-1  
Perfect score: 1274  
Sequence: 1 tttcttgtagtcttatggt.....ctcttcgtagtggaagc 1274

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9169763 seqs, 789795712 residues

Total number of hits satisfying chosen parameters: 18337526

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA, New.\*  
1: /cgn2\_6/prodata/1/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/prodata/1/pna/PCT\_NEW\_COMB.seq2.\*  
3: /cgn2\_6/prodata/1/pna/US06\_NEW\_COMB.seq.\*  
4: /cgn2\_6/prodata/1/pna/US07\_NEW\_COMB.seq.\*  
5: /cgn2\_6/prodata/1/pna/US08\_NEW\_COMB.seq.\*  
6: /cgn2\_6/prodata/1/pna/US09\_NEW\_COMB.seq.\*  
7: /cgn2\_6/prodata/1/pna/US10\_NEW\_COMB.seq.\*  
8: /cgn2\_6/prodata/1/pna/US11\_NEW\_COMB.seq.\*  
9: /cgn2\_6/prodata/1/pna/US12\_NEW\_COMB.seq.\*  
10: /cgn2\_6/prodata/1/pna/US13\_NEW\_COMB.seq.\*  
11: /cgn2\_6/prodata/1/pna/US14\_NEW\_COMB.seq.\*  
12: /cgn2\_6/prodata/1/pna/US15\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1274	100.0	3759	12	US-60-660-850-82
2	1270.8	99.7	2972	12	US-60-660-850-81
3	427.4	33.5	2240	7	US-10-955-054A-69
4	214.6	16.8	791	12	US-60-655-875-88907
5	148	11.6	480	12	US-60-655-875-96609
6	132.6	10.4	600	7	US-10-972-079-20102
7	122.6	9.6	600	7	US-10-972-079-20103
8	103.4	8.1	791	12	US-60-655-875-90124
9	103.4	8.1	809	12	US-60-655-875-97115
10	81.4	6.4	4313	8	US-10-450-763-10677
11	64.4	5.1	1552	12	US-60-655-875-69374
12	64.4	5.1	2287	7	US-60-655-875-8838
13	56.2	4.4	600	7	US-10-972-079-11909
14	55.2	4.3	14147	8	US-10-517-441-265
15	55.2	4.3	14147	8	US-10-517-441-539
16	52	4.1	17265	12	US-60-662-220-280
17	50	3.9	879	12	US-60-655-875-61962
18	50	3.9	5783	12	US-60-655-875-6546
19	48.8	3.8	17265	12	US-60-662-220-152
20	48.6	3.8	5248	12	US-60-655-875-13281
21	47.8	3.8	2030	12	US-60-655-875-30948

22	47.8	3.8	15355	8	US-10-517-441-689	Sequence 689, App
23	47.4	3.7	612	10	US-11-031-175-1357	Sequence 1357, Ap
24	46.8	3.7	1190	12	US-60-655-875-41993	Sequence 41993, A
25	46.6	3.7	1039	10	US-11-031-175-1280	Sequence 1280, Ap
26	46.4	3.6	684	12	US-60-655-875-27289	Sequence 27289, A
27	46.2	3.6	15355	8	US-10-517-441-415	Sequence 415, App
28	45.6	3.6	1038	12	US-60-655-875-15948	Sequence 15948, A
29	45.6	3.6	2138	12	US-60-655-875-21564	Sequence 21564, A
30	45	3.5	593	12	US-60-655-875-50866	Sequence 50866, A
31	45	3.5	2255	12	US-60-655-875-20345	Sequence 20345, A
32	45	3.5	3505	12	US-60-655-875-9159	Sequence 9159, Ap
33	45	3.5	5001	7	US-10-526-108-192	Sequence 192, App
34	44.8	3.5	1384	12	US-60-655-875-76917	Sequence 76917, A
35	44.8	3.5	2827	12	US-60-655-875-43121	Sequence 43121, A
36	44.6	3.5	1265	12	US-60-655-875-77202	Sequence 77202, A
37	44.6	3.5	2972	12	US-60-655-875-1982	Sequence 1982, Ap
38	44.4	3.5	942	12	US-60-655-875-30798	Sequence 30798, A
39	44.2	3.5	2058	12	US-60-655-875-27864	Sequence 27864, A
40	44	3.5	515	12	US-60-655-875-24405	Sequence 24405, A
41	44	3.5	1073	12	US-60-655-875-12406	Sequence 12406, A
42	44	3.5	5493	8	US-10-517-441-417	Sequence 417, App
43	43.8	3.4	194	12	US-60-655-875-94527	Sequence 94527, A
44	43.8	3.4	1464	12	US-60-655-875-14052	Sequence 14052, A
45	43.8	3.4	2019	12	US-60-655-875-19196	Sequence 19196, A

ALIGNMENTS

RESULT 1  
US-60-660-850-82

Sequence 82, Application US/60660850  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Ming-Yi Chiang  
APPLICANT: Lex M. Cowser  
APPLICANT: Nicholas M. Dean  
APPLICANT: Kenneth W. Dobie  
APPLICANT: Susan M. Freier  
APPLICANT: Ravi Jain  
APPLICANT: Erich Koller  
APPLICANT: Eric G. Marcussen  
APPLICANT: Robert McKay  
APPLICANT: Brett P. Monia  
APPLICANT: Donna T. Ward  
APPLICANT: Jacqueline Wyatt  
TITLE OF INVENTION: COMPOSITIONS AND THEIR USES DIRECTED TO CELL GROWTH AND/OR MAINTENANCE GENES  
FILE REFERENCE: DPTK-007905.L  
CURRENT APPLICATION NUMBER: US/60/660,850  
CURRENT FILING DATE: 2005-03-10  
NUMBER OF SEQ ID NOS: 2619  
SOFTWARE: PatentSeq version 1.0  
SEQ ID NO 82  
LENGTH: 3759  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-60-660-850-82

Query Match 100.0%; Score 1274; DB 12; Length 3759;  
Best Local Similarity 100.0%; Pred. No. 1.2e-2711;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTTCTTGTGTTGTTATGTCCTATTTCTATTTTATTTAGCATTTATTTCTATGTAGTCTAT	60
Db	1141	TTTCTTGTGTTGTTATGTCCTATTTCTATTTTATTTAGCATTTATTTCTATGTAGTCTAT	1200
QY	61	CCAAAGACGATTAAAGGAGTTCACATGTTTTCGGAACATTTTGAAGAAGAGAGCTTATC	120
Db	1201	CCAAAGACGATTAAAGGAGTTCACATGTTTTCGGAACATTTTGAAGAAGAGAGCTTATC	1260
QY	121	CAGTGTACAGATCCTAATAAGTGCACATTCAGTGTAAATTTTATTTTATTTATTTCTTT	180

Db 1261 CAGGTACAGATCCTTAATAAGTGCACATTCAGTGAATTTATTTTATATATCTTTT 1320  
Qy 181 TTAATCCTATTTTCT 240  
Db 1321 TTAATCCTATTTTCT 1380  
Qy 241 TATGGCATGTAAACATTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 300  
Db 1381 TATGGCATGTAAACATTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1440  
Qy 301 CCTTATGATTTTATTTTCAAGATGAGCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 360  
Db 1441 CCTTATGATTTTATTTTCAAGATGAGCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1500  
Qy 361 TCGAAGTGTCCACATCCACAGAGGGTGCCTGCGCTGACGTGGCAAACTGTCATCCAACT 420  
Db 1501 TCGAAGTGTCCACATCCACAGAGGGTGCCTGCGCTGACGTGGCAAACTGTCATCCAACT 1560  
Qy 421 GACTGGGAGTCTTTTGAAGTGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
Db 1561 GACTGGGAGTCTTTTGAAGTGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
Qy 481 CAATGTCAGATGGCCATCTTTTCTAGCAACTGTCGCCCAAGCTCAATGTTGTCCA 540  
Db 1621 CAATGTCAGATGGCCATCTTTTCTAGCAACTGTCGCCCAAGCTCAATGTTGTCCA 1680  
Qy 541 ACTTGGCGGGCCCTTTGGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAATCCA 600  
Db 1681 ACTTGGCGGGCCCTTTGGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAATCCA 1740  
Qy 601 GTACTTTTCCCTGTAATATGCTCTCTGGAATGGAATTAACCTGCGCACACACAGAA 660  
Db 1741 GTACTTTTCCCTGTAATATGCTCTCTGGAATGGAATTAACCTGCGCACACACAGAA 1800  
Qy 661 AAAGCAGACATCAAGAGCTCTGTAGTTTATGAGCTTATTCCTGTCGCTGCTGCTGCT 720  
Db 1801 AAAGCAGACATCAAGAGCTCTGTAGTTTATGAGCTTATTCCTGTCGCTGCTGCTGCT 1860  
Qy 721 TCTGTAAATGGCAAGCTCTCTGGATGCTGTGAATGCCCACTGATGATCAGCATAG 780  
Db 1861 TCTGTAAATGGCAAGCTCTCTGGATGCTGTGAATGCCCACTGATGATCAGCATAG 1920  
Qy 781 TCCATTAACAACCTTACAGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCT 840  
Db 1921 TCCATTAACAACCTTACAGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCT 1980  
Qy 841 GGTGCTGTGACTGGGTGATGATGCACTCTCTGTTTGGCTTTCACCTCATGTTAGTCTTA 900  
Db 1981 GGTGCTGTGACTGGGTGATGATGCACTCTCTGTTTGGCTTTCACCTCATGTTAGTCTTA 2040  
Qy 901 GAGAAACAGGAAATACGATGCTCACCAGAGTCTTCCGCAATCGTACAGCTGATAGGA 960  
Db 2041 GAGAAACAGGAAATACGATGCTCACCAGAGTCTTCCGCAATCGTACAGCTGATAGGA 2100  
Qy 961 ACAGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGTAGCTAAATGGTCAAGCGACGA 1020  
Db 2101 ACAGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGTAGCTAAATGGTCAAGCGACGA 2160  
Qy 1021 TTGACTTTGGGAAGCGACTCTCGATCTATTTCATGAAGAAATGGCAAGCCATTAATGAAT 1080  
Db 2161 TTGACTTTGGGAAGCGACTCTCGATCTATTTCATGAAGAAATGGCAAGCCATTAATGAAT 2220  
Qy 1081 AGCCAGCTGTAGTCTTTTGACACAGCAATGGCAGCTTTTGGAGAAATGGCAATTTTA 1140  
Db 2221 AGCCAGCTGTAGTCTTTTGACACAGCAATGGCAGCTTTTGGAGAAATGGCAATTTTA 2280  
Qy 1141 GGCATCAATGTAACTATTTCATGTTGTAATGGCAATCAAACTTTTCTGGCCAGTGT 1200  
Db 2281 GGCATCAATGTAACTATTTCATGTTGTAATGGCAATCAAACTTTTCTGGCCAGTGT 2340  
Qy 1201 TTAATACTTCAATTTTCAAGAAATATAGGCAACCTCTGCTGCAACCTTAAACTCTTT 1260  
Db 2341 TTAATACTTCAATTTTCAAGAAATATAGGCAACCTCTGCTGCAACCTTAAACTCTTT 2400

Qy 1261 CGGTAGGTGAAGC 1274  
Db 2401 CGGTAGGTGAAGC 2414

RESULT 2  
US-60-660-850-81  
; Sequence 81, Application US/6060850  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Ming-Yi Chiang  
; APPLICANT: Lex M. Cowser  
; APPLICANT: Nicholas M. Dean  
; APPLICANT: Kenneth W. Dobie  
; APPLICANT: Susan M. Freier  
; APPLICANT: Ravi Jain  
; APPLICANT: Eric G. Marcusson  
; APPLICANT: Robert McKay  
; APPLICANT: Brett P. Monia  
; APPLICANT: Donna T. Ward  
; APPLICANT: Jacqueline Wyatt  
; TITLE OF INVENTION: COMPOSITIONS AND THEIR USES DIRECTED TO CELL GROWTH AND/OR  
; FILE REFERENCE: DPTK-0079US.L  
; CURRENT APPLICATION NUMBER: US/60/660,850  
; CURRENT FILING DATE: 2005-03-10  
; NUMBER OF SEQ ID NOS: 2619  
; SOFTWARE: PatentSeq version 1.0  
; SEQ ID NO 81  
; LENGTH: 2972  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-60-660-850-81

Query Match 99.7%; Score 1270.8; DB 12; Length 2972;  
Best Local Similarity 99.8%; Pred. No. 6.1e-271;  
Matches 1272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TTTCTTTAGTTGTTTATGTTCCATTTTCTATTATAGCAATTTATTTCTATCTATCTATCTAT 60  
Db 741 TTCTTTTATGTTTATGTTCCATTTTCTATTATAGCAATTTATTTTATTTATTTATTTATTT 800  
Qy 61 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAAATTTTGAAGAGAGAGCTTATC 120  
Db 801 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAAATTTTGAAGAGAGAGCTTATC 860  
Qy 121 CAGGTACAGATCCTTAATAAGTGCACATTCAGTCTAAATTTATTTTATTTATTTATTTAT 180  
Db 861 CAGGTACAGATCCTTAATAAGTGCACATTCAGTCTAAATTTATTTTATTTATTTATTTAT 920  
Qy 181 TTAATCCTATTTTCT 240  
Db 921 TTAATCCTATTTTCT 980  
Qy 241 TATGGCATGTAAACATTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 300  
Db 981 TATGGCATGTAAACATTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1040  
Qy 301 CCTTATGATTTTATTTTCAAGATGAGCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 360  
Db 1041 CCTTATGATTTTATTTTCAAGATGAGCGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1100  
Qy 361 TCGAAGTGTCCACATCCACAGAGGGTGCCTGCGCTGACGTGGCAAACTGTCATCCAACT 420  
Db 1101 TCGAAGTGTCCACATCCACAGAGGGTGCCTGCGCTGACGTGGCAAACTGTCATCCAACT 1160  
Qy 421 GACTGGGAGTCTTTTGAAGTGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
Db 1161 GACTGGGAGTCTTTTGAAGTGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220  
Qy 481 CAATGTCAGATGGCCATCTTTTCTAGCAACTGTCGCCCAAGCTCAATGTTGTCCA 540

Db 1221 CAATGTGAGAGTGGCAATCTTTGTTGAGCACTGTGCGCCAAAGCTCAATGTTGTCCA 1280  
QY 541 ACTTGGCGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATTTCA 600  
Db 1281 ACTTGGCGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATTTCA 1340  
QY 601 GTACTTTCCCTGTAATAATGCTCTTCGGATGTGAATAAATCTGCGCACACAGAA 660  
Db 1341 GTACTTTCCCTGTAATAATGCTCTTCGGATGTGAATAAATCTGCGCACACAGAA 1400  
QY 661 AAAGCAGACCATGAAGAGCTCTGTAGTTTGGCTTATTCCTGTCGTCGCCCTGGTGCT 720  
Db 1401 AAAGCAGACCATGAAGAGCTCTGTAGTTTGGCTTATTCCTGTCGTCGCCCTGGTGCT 1460  
QY 721 TCTGTAAATGGCAAGCTCTGTAGTTTGGCTTATTCCTGTCGTCGCCCTGGTGCT 780  
Db 1461 TCTGTAAATGGCAAGCTCTGTAGTTTGGCTTATTCCTGTCGTCGCCCTGGTGCT 1520  
QY 781 TCCATTACACCCCTACAGGAGAGGATATAGTTTTCCTGCTACAGACATTAATCTTCCT 840  
Db 1521 TCCATTACACCCCTACAGGAGAGGATATAGTTTTCCTGCTACAGACATTAATCTTCCT 1580  
QY 841 GGTGCTGTGACTGGGTGATGATGCGAGTCTGCTGTTTGGCTTTCACCTTCACTGTTAGTCTTA 900  
Db 1581 GGTGCTGTGACTGGGTGATGATGCGAGTCTGCTGTTTGGCTTTCACCTTCACTGTTAGTCTTA 1640  
QY 901 GAGAAACAGGAAATAATACGATGCTCACCAGCAGTCTTTCGCAATCGTACAGCTGTATGGA 960  
Db 1641 GAGAAACAGGAAATAATACGATGCTCACCAGCAGTCTTTCGCAATCGTACAGCTGTATGGA 1700  
QY 961 ACACGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGGCTTTCGCAATCGTACAGCTGTATGGA 1020  
Db 1701 ACACGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGGCTTTCGCAATCGTACAGCTGTATGGA 1760  
QY 1021 TTGACTTGGAGCGAGCTCTCGATCTATTCATGAAGGAAATGCAACAGCATTATGAAT 1080  
Db 1761 TTGACTTGGAGCGAGCTCTCGATCTATTCATGAAGGAAATGCAACAGCATTATGAAT 1820  
QY 1081 AGCGACTGTCTAGCTTTGACACAGCAGTATGCAAGCTTTTTCGCAAAATGCAATTTTA 1140  
Db 1821 AGCGACTGTCTAGCTTTGACACAGCAGTATGCAAGCTTTTTCGCAAAATGCAATTTTA 1880  
QY 1141 GGCAATCAATGAACTATTTCCATGTTGAAATGGCAATCAAAATTTTTCGCCAGTGT 1200  
Db 1881 GGCAATCAATGAACTATTTCCATGTTGAAATGGCAATCAAAATTTTTCGCCAGTGT 1940  
QY 1201 TTAATACTTCAGTTTTCACAGAAATAGGCCATCTGCTGCCAACCTTAAACTCTTT 1260  
Db 1941 TTAATACTTCAGTTTTCACAGAAATAGGCCATCTGCTGCCAACCTTAAACTCTTT 2000  
QY 1261 CGGTAGGTGGAAGC 1274  
Db 2001 CGGTAGGTGGAAGC 2014

RESULT 3

US-10-955-054A-69

; Sequence 69, Application US/10955054A

; GENERAL INFORMATION:

; APPLICANT: PUSZTAI, LAJOS

; APPLICANT: SYMMANS, W. FRASER

; APPLICANT: HESS, KENNETH R.

; APPLICANT: AYERS, MARK

; APPLICANT: SPEC. JAMES

; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY

; FILE REFERENCE: UTX:880US

; CURRENT APPLICATION NUMBER: US/10/955,054A

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 195

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 69

; LENGTH: 2240

; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-955-054A-69  
Query Match 33.5%; Score 427.4; DB 7; Length 2240;  
Best Local Similarity 73.7%; Pred. No. 5.7e-85;  
Matches 558; Conservative 0; Mismatches 196; Indels 3; Gaps 1;  
QY 416 ACAATGACTTTGGGAGTCTTTTGGAGTCCAGTCTGCTTTGACTATGTGTACCGCCCAACACTCAGATGTT 535  
Db 738 ACCACGAGCTGACCTCGCTCTTTCGAGTCCGCTCTGCTTTGACTATGTCTCTGCTCCCTA 797  
QY 476 TTCTTCAATGTGAGAGTGGCCATCTTGTGTTGAGCACTGTCGCCCAACACTCAGATGTT 535  
Db 798 TTCTGAGTGGCCAGGCGGGCACCTGGTGTGTAAACAATGCCGCCAGAAAGTTGAGCTGT 857  
QY 536 GTCCAACTTGGCGGGCCCTTTGGGATC---CATTCGCAACTTTGGCTATGGAGAAAGTGG 592  
Db 858 GCCGAGCTGCGAGGGGGCCCTGACGCCAGCAGCATCAGGAACCTGGCTATGGAGAAAGTGG 917  
QY 593 CTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTTCGATGTGAAATAAATCTCTGCCAC 652  
Db 918 CTTGGCAGTCTCTGTTTCCCTGTAGTATGCCACACGGGCTGTTCCCTGACCTGCACC 977  
QY 653 ACACAGAAAAGCAGACCATGAGAGCTCTGTAGTGTAGGCTTATTCCTGCTCGTGCC 712  
Db 978 ATACGGAGAAACAGAAACATGTAAGACATATGTAATACCGCCCTACTCTCTGCCCATGTC 1037  
QY 713 CTGTGCTCTTCTGTAAATGCGCAAGGCTCTCTGATGCTGTAAATGCCCATCTCATGATC 772  
Db 1038 CTGTGCTCTTCTGTAAATGCGCAAGGCTCTCTGATGCTGTAAATGCCCATCTCATGATC 1097  
QY 773 AGCATAAGTCCATTACAACTTACAGGGAGAGGATATAGTCTTTTCTTGTACAGACATTA 832  
Db 1098 CCCACAAGAGCATTACCACTTTCAGGAAGAGACATCGTCTTTCTAGCTACAGACATTA 1157  
QY 833 ATCTTCTGCTGCTGTTGATGCTGGGTGATGATGAGTCTCTGTTTGGCTTTCATCTTATGT 892  
Db 1158 ACTTGCAGGGGCTGTGCACTGGGTGATGATGAGTCTGTTTGGCCATCACTTTCATGTC 1217  
QY 893 TAGTCTTACAGAAACAGGAAATAACGATGCTCACCAGCAGTCTTTCGCAATCGTACAGC 952  
Db 1218 TGTGCTGAGAAACAGGAAGTACGAAGGCCACAGCAGTCTTTCGCAATCGTCTGTC 1277  
QY 953 TGATAGGAAACAGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGGAGCTTAAATGCTCAT 1012  
Db 1278 TCATTGGCACCCGCAAGCAAGCGGAGAACTTTTGGCTTACAGACTGGAGTGAATGGGAAC 1337  
QY 1013 GCGCAGATGATGCTGGGAGGCACTCTCTGATCTATTTCATGAGGAATTCGAACAGCCA 1072  
Db 1338 GCGGAGATGATGCTGGGAGGCACTCTCTGATCTATTTCATGAGCTGCTGCTGCGGCCA 1397  
QY 1073 TTATGAATAGCAGTCTGATGCTTTGACACAGCAATTCACAGCTTTTTCGAGAAATG 1132  
Db 1398 TCATGAACAGCAGCTGCTGTTTTCGACAGCCATACATCTTTTTCGAGAAATG 1457  
QY 1133 GCAATTTAGGCAATCAATGTAATTTTCCATGTTG 1169  
Db 1458 GGAACCTTGAATCAATGTTACTATTTCTACATGTTG 1494

RESULT 4

US-60-655-875-88907

; Sequence 88907, Application US/60655875

; GENERAL INFORMATION:

; APPLICANT: Boukharov, Andrey

; APPLICANT: Du, Zijing

; APPLICANT: Guo, Liang

; APPLICANT: Kovalic, David

; APPLICANT: Lu, Maolong

; APPLICANT: McGarter, James

; APPLICANT: Miller, Nancy

; APPLICANT: Williams, Deryck

APPLICANT: Vaudin, Mark  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF  
FILE REFERENCE: 38-21(53885)  
CURRENT APPLICATION NUMBER: US/60/655,875  
CURRENT FILING DATE: 2005-02-24  
NUMBER OF SEQ ID NOS: 171306  
SEQ ID NO 88907  
LENGTH: 791  
TYPE: DNA  
ORGANISM: Heterodera glycines  
FEATURE:  
OTHER INFORMATION: SCN-specific regions: Sequence locations=1-791  
US-60-655-875-88907

Query Match 16.8%; Score 214.6; DB 12; Length 791;  
Best Local Similarity 63.9%; Pred. No. 4e-38;  
Matches 325; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 412 TCACAAATGACTTGGCGAGTCTTTTGGAGTCCAGTCTGCTTTGACTATGTTAGCG 471  
Db 258 TCAAAATCTGAATTTCTGTGAGCTTCGAATGTCGGTTTGTCTTATTACATGCTTCG 317  
QY 472 CCCATTCTTCAATGTCAGAGTGGCCATCTTTTGTAGCAACTGTGCGCCAAAGCTCACA 531  
Db 318 CCATATCTGCAATGTCATCGGTCATCTGTTTGTGGAATTTGCTTCTTAAGTCACT 377  
QY 532 TGTGTTCCAACTTGGCGGGGCCCTTTGGGATCAATTCGCAACTTGGCTATGAGAAAGTG 591  
Db 378 GCTTGGCCGACATGTCGAGTCCCAATTCGAGCATGAAGAAATCTTGGAAATGGAGAAAT 437  
QY 592 GCTAATTCAGTACTTTCCCTCTAATATGCTCTTCTGGATGTAATAACTCTGCGCA 651  
Db 438 GCAGCAAGCCTGCAATTTCCATGCAATTTGCTTCTGAGTGGTGCACCCCAATCTTTTAT 497  
QY 652 CACACAGAAAAGCAGACCAATGAGAGCTCTGTGAGTTTAGGCTTTATTCCTGTCGCTGC 711  
Db 498 CATCATGAGAAATGAGACGAGAGATTTGGAGTTTCGCTTACCAATGCCCTTGT 557  
QY 712 CTTGGTCTTCTGTAATGGCAAGGCTCTCGATGCTGTATGCTCCCATCTGATGCAAT 771  
Db 558 CTTGGAGCTTCTGCAATGCAATGCAAGGAGTCTGGCAGATGTGATGAGCCATCTAATGAAA 617  
QY 772 CAGCATAGTCCATTACAACCTACAGGAGAGATATAGTTTCTTGTCTACAGACATT 831  
Db 618 TTGCAAAATCAATCAACCTTTGCAAGGCAAGACATTTGTTTCTGCAACTGACATC 677  
QY 832 AATCTTCTGCTGCTGTGACTGGGTGATGATGATGCTCTGTTTGGCTTTCACTTCATG 891  
Db 678 AACCTCTGCTGCTGCTGATTTGGTAAATGATGCAATCGTGTACAAAGCCTTTTTCATG 737  
QY 892 TTAGTCTTAGAGAAACAGGAAATACGA 920  
Db 738 CTCGCTCTGGAGGCAAGCAACCAAGA 766

RESULT 5  
US-60-655-875-96609  
Sequence 96609, Application US/60655875  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey  
APPLICANT: Du, Zijing  
APPLICANT: Guo, Liang  
APPLICANT: Kovalic, David  
APPLICANT: Lu, Maolong  
APPLICANT: McCarter, James  
APPLICANT: Miller, Nancy  
APPLICANT: Williams, Deryck  
APPLICANT: Vaudin, Mark  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF

FILE REFERENCE: 38-21(53885)  
CURRENT APPLICATION NUMBER: US/60/655,875  
CURRENT FILING DATE: 2005-02-24  
NUMBER OF SEQ ID NOS: 171306  
SEQ ID NO 96609  
LENGTH: 481  
TYPE: DNA  
ORGANISM: Heterodera glycines  
FEATURE:  
OTHER INFORMATION: SCN-specific regions: Sequence locations=1-481  
US-60-655-875-96609

Query Match 11.6%; Score 148; DB 12; Length 481;  
Best Local Similarity 61.8%; Pred. No. 1.8e-23;  
Matches 235; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 412 TCACAAATGACTTGGCGAGTCTTTTGGAGTCCAGTCTGCTTTGACTATGTTAGCG 471  
Db 99 TCAAAATCTGAATTTCTGTGAGCTTCGAATGTCGGTTTGTCTTATTACATGCTTCG 158  
QY 472 CCCATTCTTCAATGTCAGAGTGGCCATCTTTTGTAGCAACTGTGCGCCAAAGCTCACA 531  
Db 159 CCATATCTGCAATGTCATCGGTCATCTGTTTGTGGAATTTGCTTCTTAAGTCACT 218  
QY 532 TGTGTTCCAACTTGGCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTG 591  
Db 219 GCTTGGCCGACATGTCGAGTCCCAATTCGAGCATGAAGAAATCTTGGAAATGGAGAAAT 278  
QY 592 GCTAATTCAGTACTTTCCCTCTAATATGCTCTTCTGGATGTAATAACTCTGCGCA 651  
Db 279 GCAGCAAGCCTGCAATTTCCATGCAATTTGCTTCTGAGTGGTGCACCCCAATCTTTTAT 338  
QY 652 CACACAGAAAAGCAGACCAATGAGAGCTCTGTGAGTTTAGGCTTTATTCCTGCTGCTGC 711  
Db 339 CATCATGAGAAATGAGACGAGAGATTTGGAGTTTCGCTTACCAATGCCCTTGT 398  
QY 712 CTTGGTCTTCTGTAATGGCAAGGCTCTCGATGCTGTATGCTCCCATCTGATGCAAT 771  
Db 399 CTTGGAGCTTCTGCAATGCAATGCAAGGAGTCTGGCAGATGTGATGAGCCATCTAATGAAA 458  
QY 772 CAGCATAGTCCATTACAAC 791  
Db 459 TTGCAAAATCAATCAAC 478

RESULT 6  
US-10-972-079-20102/c  
Sequence 20102, Application US/10972079  
GENERAL INFORMATION:  
APPLICANT: MMI GENOMICS, INC.  
APPLICANT: Denise, Sue K.  
APPLICANT: ROSENFELD, David  
APPLICANT: KERR, Richard  
APPLICANT: BATES, Stephen  
APPLICANT: HOLM, Tom  
TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEEF LIVESTOCK  
FILE REFERENCE: MM11110-2  
CURRENT APPLICATION NUMBER: US/10/972,079  
CURRENT FILING DATE: 2004-10-22  
PRIOR APPLICATION NUMBER: US 60/514,333  
PRIOR FILING DATE: 2003-10-24  
NUMBER OF SEQ ID NOS: 96631  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 20102  
LENGTH: 600  
TYPE: DNA  
ORGANISM: Chicken 19866894230253\_1  
US-10-972-079-20102

Query Match 10.4%; Score 132.6; DB 7; Length 600;  
Best Local Similarity 67.6%; Pred. No. 4.6e-20;  
Matches 186; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

US-60-655-875-90124/c  
; Sequence 90124, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijing  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 90124  
; LENGTH: 791  
; TYPE: DNA  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; OTHER INFORMATION: SCN-specific regions: Sequence locations=1-21,43-791  
US-60-655-875-90124

Query Match 8.1%; Score 103.4; DB 12; Length 791;  
Best Local Similarity 65.2%; Pred. No. 1.3e-13;  
Matches 152; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 926 ACCAGCAGTTCCTCGCAATCGTACAGCTGATAGGACACAGCAAGCAAGCTGAAAAATTTTG 985  
Db 782 ACCAAATGTTCTACGCGAGTTGTTCAACTAATTTGGGACCAAAAGGAAGCCGAAAAATTCG 723

Qy 986 CTTACCGACTTGAGCTAAATGGTCATAGGCGAGATGACTTGGGAAGCGACTCTCTCGAT 1045  
Db 722 TTTCACAAATGGGAATTTATCCAAATCGGCGTCGACTTTGCTGGGAAGCCAGTCCACGAA 663

Qy 1046 CTATTCATGAAGGAAATTTGCAACAGCCATTATGAATAGCGACTCTCTAGTCTTTTGACACCA 1105  
Db 662 GTATTACAGAGGGGTGCTGCGCGCAATTTCTCAATCAGACTGTTTGGCTTTTGACACAA 603

Qy 1106 GCATTGCACAGCTTTTTCGAGAAAAATGGCAATTTAGGCATCAATGTAACCTATT 1158  
Db 602 ATCATGCGAATTTTTTGTCTGAAACCGGAAATTTGGGAATAAATGTACTATT 550

RESULT 9  
US-60-655-875-97115/c  
; Sequence 97115, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijing  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 97115  
; LENGTH: 809  
; TYPE: DNA  
; ORGANISM: Heterodera glycines  
; FEATURE:



```
; OTHER INFORMATION: SCN-specific regions: Sequence locations=1-809
US-60-655-875-97115

Query Match      8.1%; Score 103.4; DB 12; Length 809;
Best Local Similarity 65.2%; Pred. No. 1.3e-13;
Matches 152; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 926 ACCACGAGTTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAAATTTTG 985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 800 ACCAAATGTTCTACGAGTTGTTCACTAAATTTGGACCAAAAGGAAGCCGAAATTTTCG 741

QY 986 CTTACCGACTTGAGCTAAATGTCATAGCGGACGATTCGACTTGGGAAGCGACTCTCTCGAT 1045
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 740 TTTCAAATTTGGAATATCCAAATTCGCGTCTGACTTTGCTGGGAAGCCAGTCCACGAA 681

QY 1046 CTTATTCATGAAGGAATTGCAACAGCCATTATGAATAGGAGCTGTCTAGCTTTTGACACA 1105
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 680 GTATTACGAGGGGTGCGTCTGCGCAATTTCTCAATCAGACTGTTTGGCTTTTGACACAA 621

QY 1106 GCAATGCAAGCTTTTTCAGAAAATGCAATTTAGGCATCAATGTAAGTATT 1158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 620 ATCATGCGAAATTTTCTGCTGAAAACGGGAATTTGGGAATAAATGTACTATT 568

RESULT 10
US-10-450-763-10677/c
; Sequence 10677, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 10677
; LENGTH: 4313
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SIMILAR
; LOCATION: (13)...(519)
; OTHER INFORMATION: 75% homologous to Homo sapiens Spi40 protein, accession number
; OTHER INFORMATION: U63420, Smith-Waterman Score=622.
US-10-450-763-10677

Query Match      6.4%; Score 81.4; DB 8; Length 4313;
Best Local Similarity 64.9%; Pred. No. 1.3e-08;
Matches 137; Conservative 0; Mismatches 71; Indels 3; Gaps 1;

QY 416 ACAATGACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCCA 475
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2411 ACCACGAGCTGACTCGCTCTTCGAGTGTCCGCTGTGCTTTGACTATGTCTGCTCTCTTA 2352

QY 476 TTCTTCAATGTCAAGTGGCCATCTTTGTTGTAGCAACTGTGCCCCAAAGCTCACATGTT 535
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2351 TTCTGCAATGCCAGCGCGGCACCTGTGTGTAAACCAATGCCGCCAGAAAGTTGAGCTGT 2292

QY 536 GTCCAACTTGGCGGGGCCCTTTGGATC---CATTCGCACTTGGCTATGGAGAAATGG 592
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2291 GCCCGAGCTGCGAGGGGGCCCTGACGCCCCAGCATCAGGAACCAAGATGAGAGAGTGG 2232

QY 593 CTAATTCAGTACTTTTCCCTGTAAATATGC 623
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2231 CCTCGGAGTCTCTGTTTCCCGGCCCAATGC 2201

; OTHER INFORMATION: Contig location: Contig_ID=SeqID_8838; Location=94...1645;
US-60-655-875-69374
; Sequence 69374, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETEROERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 69374
; LENGTH: 1552
; TYPE: DNA
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Genomic contig location:
; OTHER INFORMATION: Strand=+
; FEATURE:
; OTHER INFORMATION: SCN-specific regions: Sequence locations=1-380,402-903,925-1120,11
; OTHER INFORMATION: -1303,1328-1408,1449-1552
; FEATURE:
; OTHER INFORMATION: Gene Ontology: Mol. function=transcription co-repressor activity;
; OTHER INFORMATION: Cellular component=cytoplasm; Biological process=small GTPase
; OTHER INFORMATION: mediated signal transduction
US-60-655-875-69374

Query Match      5.1%; Score 64.4; DB 12; Length 1552;
Best Local Similarity 62.3%; Pred. No. 6.1e-05;
Matches 101; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 651 ACACACAGAAAAGCAGACCACTCTGTGAGTTTAGCCCTTATCTCTGTCCTG 710
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 734 ATACTCAATATATTGGAAATTTACTTGTCTTACTAGTCTTACCAATGCCCTTG 793

QY 711 CCTGTGCTTCTCTGTAATGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATCA 770
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 TCCTGGAGCTTCTGCAATGCAAGGAGTCTGGCAGATGTGATGAGCATCTAATGAA 853

QY 771 TCAGCATAGTCCATTACAACTTACAGGAGGATAGT 812
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 ATTGCAAAATCAATCACAACCTTTGCAAGCAATTTTAAAT 895

RESULT 12
US-60-655-875-8838
; Sequence 8838, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETEROERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
```



```

; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARK, Almuth
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 539
; LENGTH: 14147
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-539

Query Match      4.3%; Score 55.2; DB 8; Length 14147;
Best Local Similarity 48.4%; Pred. No. 0.0097;
Matches 153; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 1 TTCTTTAGTGTGTTAGTGCATTTCTATTTAGCATTTATTTCTATGACTCTAT 60
Db 10843 TTTGTTTTTTTTTGTGTTGTTTTTTTTTTTTTTTGAATTTATGTTAGTTTTTTTTTT 10902

QY 61 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTGAAAGAGAGCTTATC 120
Db 10903 TAGTGTATTGTTAGGTATTTTATGTTTGGGTTTTTATATGTTGTTGTTTATT 10962

QY 121 CAGGTACAGATCCTAATAAAGTGCACATTCAGTGTAAATTTATTTTAAATATCTTT 180
Db 10963 TGGTTGTTGTTTATAGATATTTATGAGTTAGTTTTTTTATTTTGGTTTGGTTTGT 11022

QY 181 TTAATCCTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 11023 TAGTGTGTTTTTTTTTGGTGTATTTTGTGTTTTTTTGTGTTTTTGTGTTTTT 11082

QY 241 TAGGCATGTAAACATTTATTTAATAAGTACATGTTATGTTATTTTCTCTCTCT 300
Db 11083 TTTTGTGTTGTTTTTTTTTATTAATTTTATTTATTTATTTGATGATTTTGTATATGTT 11142

QY 301 CCTTATGTTATTTT 316
Db 11143 TGTTTTTTTTTTATT 11158

```

Search completed: April 25, 2005, 00:35:39  
Job time : 478 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 22:53:26 ; Search time 758 Seconds  
(without alignments)  
9949.539 Million cell updates/sec

Title: US-10-679-246-1  
Perfect score: 1274  
Sequence: 1 ttctttagttgtttatgt.....cttttcgtagtggaagc 1274

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 15

Total number of hits satisfying chosen parameters: 17572

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N\_Geneseq\_16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1274	100.0	1274	4	Aac67281 Human Sia
2	1274	100.0	2924	5	Abv28804 Human pro
3	1274	100.0	2924	5	Abv25468 Human pro
4	1274	100.0	2924	5	Abv22970 Human pro
5	1274	100.0	6107	4	Aal05382 Human rep
6	1274	100.0	6107	4	Ab198246 Human tes
7	1223	96.0	2829	11	Adm01584 Human cdn
8	1223	96.0	2829	13	AdS34362
9	1104	86.7	2128	13	AdS34361
10	1041	81.7	2440	11	AdI31262 Human cdn
11	958	75.2	1540	13	AdS34365
12	958	75.2	2048	13	AdS34360
13	957	75.1	2454	13	AdS34359
14	947	74.3	1886	13	AdS34363
15	881	69.2	4090	10	AdF81981
16	838	65.8	2034	13	AdS34364
17	780	61.2	1884	2	Aat64820 Tumour su
18	581	45.6	3128	10	AdI40343 Human pur
19	356	27.9	1535	3	Aac98856 Human pan
20	203	15.9	466	4	Aai87326 Human pol

Abv38421	Human pro	405	12.5	159	C	21
Abv08519	Human pro	277	12.1	154	C	22
Abv43789	Human pro	417	10.9	139	C	23
Abv48461	Human pro	473	10.0	128	C	24
Abv13828	Human pro	411	9.3	118	C	25
Abv34940	Human pro	441	9.3	118	C	26
Abv04659	Human pro	438	6.7	85	C	27
Abi99429	Mouse isc	1968	6.0	77	C	28
Abx44846	Bovine ES	360	4.9	62	C	29
Abv18678	Human pro	426	4.1	52	C	30
Aal01548	Human rep	222	2.6	33	C	31
Ab196986	Human tes	222	2.6	33	C	32
Abi99430	Mouse isc	2198	2.4	30	C	33
AdF54673	Human SIA	33	1.9	24	C	34
AbS69096	Novel mur	226	1.8	23	C	35
AdF54668	Human SIA	21	1.6	21	C	36
AdF54669	Human SIA	21	1.6	21	C	37
Ach28397	Human adu	442	1.6	21	C	38
Ach28129	Human adu	477	1.6	21	C	39
AdI14958	Human gla	1094	1.6	21	C	40
AdI14957	Human gla	1150	1.6	21	C	41
Adel3891	Human opt	46951	1.6	21	C	42
Aac67288	Human SIA	28	1.6	20	C	43
AcS55739	Human sig	257	1.6	20	C	44
AdI55595	Human pol	257	1.6	20	C	45
Ab108593	Drosophil	834	1.6	20	C	46
Adr25256	Breast ca	975	1.6	20	C	47
Acc50292	Breast ca	2240	1.6	20	C	48
Adp13401	Renal cel	2240	1.6	20	C	49
Adr14122	Human NF-	2240	1.6	20	C	50
Adp54830	Human PRO	2240	1.6	20	C	51
Acc50293	Breast ca	2502	1.6	20	C	52
Ab119804	Drosophil	2961	1.6	20	C	53
Ab108592	Drosophil	3157	1.6	20	C	54
Aag75209	ALI-1 (ac	8342	1.6	20	C	55
Aat6333	MLL gene	8391	1.6	20	C	56
Aag3478	MLL gene	8392	1.6	20	C	57
Aaa99882	Human MLL	8392	1.6	20	C	58
Continuation (10 o		110000	1.6	20	C	59
Aac67289	Human SIA	19	1.5	20	C	60
Abv01160	Human pro	231	1.5	19	C	61
Abd11499	Pseudomon	345	1.5	19	C	62
Abv40469	Human pro	447	1.5	19	C	63
Abv31500	Human pro	447	1.5	19	C	64
Abv47042	Human pro	470	1.5	19	C	65
AdQ56323	Novel can	547	1.5	19	C	66
Ach79377	Human gn	564	1.5	19	C	67
Acn62684	Cotton de	567	1.5	19	C	68
Aal04022	Human rep	605	1.5	19	C	69
Aal04020	Human rep	605	1.5	19	C	70
Aas40419	DNA encod	605	1.5	19	C	71
Aas40417	DNA encod	605	1.5	19	C	72
Adj09623	Human pro	605	1.5	19	C	73
Adj09625	Human pro	605	1.5	19	C	74
Abd11398	Pseudomon	636	1.5	19	C	75
Abv10329	Human pro	850	1.5	19	C	76
Adc79411	Human G p	1014	1.5	19	C	77
Abq88091	Human GPC	1015	1.5	19	C	78
Aag87727	Human aux	1473	1.5	19	C	79
Aat28393	Human cyc	1473	1.5	19	C	80
AdQ35351	Novel mou	1662	1.5	19	C	81
Adi12649	Human ENZ	1795	1.5	19	C	82
ACA34434	Prokaryot	1806	1.5	19	C	83
Aas00019	Pseudomon	1884	1.5	19	C	84
Aas00017	Pseudomon	1884	1.5	19	C	85
Aas00018	Pseudomon	1884	1.5	19	C	86
AAL40123	Isoprenoi	1884	1.5	19	C	87
Abd11213	Pseudomon	1920	1.5	19	C	88
Ab107583	Drosophil	1951	1.5	19	C	89
Adi31729	Human cdn	1995	1.5	19	C	90
Adq39097	Human SNP	1995	1.5	19	C	91
Ada71685	Rice gene	2009	1.5	19	C	92
Aat11379	Cytochrom	2009	1.5	19	C	93

94	19	1.5	2009	2	AAV441155	AAv441155 Human cyt	18	1.4	498	2	AAT63116	Aat63116 Human sol
95	19	1.5	2188	13	ACN41238	Acn41238 Human dia	18	1.4	501	2	AAT63112	Aat63112 Human sol
96	19	1.5	2191	13	ACN41233	Acn41233 Human dia	18	1.4	504	2	AAQ36621	Aaq36621 Soluble K
97	19	1.5	2239	13	ACN41232	Acn41232 Human dia	18	1.4	506	2	AAV07350	Aav07350 Stem cell
98	19	1.5	2258	2	AAV11380	Aat11380 Cytochrom	18	1.4	506	2	AAV07364	Aav07364 Stem cell
99	19	1.5	2258	2	AAV44158	AAv44158 Human cyt	18	1.4	506	2	AAV07349	Aav07349 Stem cell
100	19	1.5	2258	6	ABN95871	Abn95871 Gene #236	18	1.4	506	2	AAV07348	Aav07348 Stem cell
101	19	1.5	2258	11	AD131669	Adi31669 Human cDN	18	1.4	506	2	AAV07360	Aav07360 Stem cell
102	19	1.5	2268	6	ABS65367	Abse65367 Clostridi	18	1.4	506	2	AAV07365	Aav07365 Stem cell
103	19	1.5	2271	4	AA165841	Aa165841 Nucleotid	18	1.4	506	2	AAV07359	Aav07359 Stem cell
104	19	1.5	2271	6	ABS65378	Abse65378 Clostridi	18	1.4	506	2	AAV07351	Aav07351 Stem cell
105	19	1.5	2297	13	ACN41236	Acn41236 Human dia	18	1.4	506	2	AAV07353	Aav07353 Stem cell
106	19	1.5	2341	13	ACN41235	Acn41235 Human dia	18	1.4	506	2	AAV07347	Aav07347 Stem cell
107	19	1.5	2349	13	ACN41234	Acn41234 Human dia	18	1.4	506	2	AAV07354	Aav07354 Stem cell
108	19	1.5	2389	13	ACN41231	Acn41231 Human dia	18	1.4	506	2	AAV07352	Aav07352 Stem cell
109	19	1.5	2502	10	ADD29610	Add29610 Human tum	18	1.4	506	2	AAV07363	Aav07363 Stem cell
110	19	1.5	3001	3	AAH51808	Aah51808 Chromosom	18	1.4	506	2	AAV07366	Aav07366 Stem cell
111	19	1.5	3127	4	ABL09644	Ab109644 Drosophil	18	1.4	506	2	AAV07367	Aav07367 Stem cell
112	19	1.5	3662	13	ADT05448	Adt05448 Haemophil	18	1.4	506	2	AAV07346	Aav07346 Stem cell
113	19	1.5	3991	4	ABL07926	Ab107926 Drosophil	18	1.4	506	2	AAV07357	Aav07357 Stem cell
114	19	1.5	4026	4	ABL07582	Ab107582 Drosophil	18	1.4	506	2	AAV07344	Aav07344 Stem cell
115	19	1.5	4372	4	ABL07954	Ab107954 Drosophil	18	1.4	506	2	AAV07355	Aav07355 Stem cell
116	19	1.5	4515	11	ABD11433	Abd11433 Pseudomon	18	1.4	506	2	AAV07356	Aav07356 Stem cell
117	19	1.5	4821	13	ADT05416	Adt05416 Haemophil	18	1.4	506	2	AAV07345	Aav07345 Stem cell
118	19	1.5	6131	6	ABQ67024	Abq67024 Human ang	18	1.4	506	2	AAV07361	Aav07361 Stem cell
119	19	1.5	6621	3	AAV70188	Aaa70188 Plasmodiu	18	1.4	506	2	AAV07362	Aav07362 Stem cell
120	19	1.5	6641	6	ABN80003	Abn80003 Human che	18	1.4	507	2	AAV07336	Aav07336 Stem cell
121	19	1.5	7597	6	ABL33013	Ab133013 Human imm	18	1.4	507	2	AAV07334	Aav07334 Stem cell
122	19	1.5	7728	6	ABL32077	Ab132077 Human imm	18	1.4	507	2	AAV07335	Aav07335 Stem cell
123	19	1.5	7728	6	ABK40072	Abk40072 Human che	18	1.4	507	2	AAV07339	Aav07339 Stem cell
124	19	1.5	9983	6	ABK32116	Ab132116 Human imm	18	1.4	507	2	AAV07337	Aav07337 Stem cell
125	19	1.5	13249	6	ABK31176	Abk31176 Signal tr	18	1.4	507	2	AAV07333	Aav07333 Stem cell
126	19	1.5	13249	6	ABL70131	Ab170131 Chemical	18	1.4	507	2	AAV07338	Aav07338 Stem cell
127	19	1.5	17211	6	ABL32654	Ab132654 Human imm	18	1.4	507	2	AAV07340	Aav07340 Stem cell
128	19	1.5	17211	6	ABL32654	Ab132654 Human imm	18	1.4	507	2	AAV07341	Aav07341 Stem cell
129	19	1.5	17721	5	ABR33729	Ab133729 Human imm	18	1.4	512	2	AAV073617	Aat73617 Human ste
130	19	1.5	20320	5	ABA21464	Ab21464 Human ner	18	1.4	512	2	AAV073616	Aat73616 Human ste
131	19	1.5	32187	5	AAS32249	Ab32249 Human DNA	18	1.4	512	2	AAV073613	Aat73613 EST clone
132	19	1.5	32187	6	ABS87552	Ab87552 Novel hum	18	1.4	512	2	AAV899273	Aav89273 EST clone
133	19	1.5	33053	6	ABQ67006	Abq67006 Human ang	18	1.4	522	12	ACH76004	Ach76004 Human gen
134	19	1.5	43991	12	ADJ12398	Adj12398 DNA fragm	18	1.4	555	5	AAF68445	Aaf68445 Human lun
135	19	1.5	58822	9	ADA02540	Ada02540 Human TCO	18	1.4	555	8	ACA10685	Abk38356 cDNA enco
136	19	1.5	58822	10	ADB72278	Adb72278 Human TCO	18	1.4	555	8	ACA10685	ACA10685 Human lun
137	19	1.5	58822	10	ADE95788	Ad95788 Human TCO	18	1.4	555	8	ABX99636	Abx99636 Lung canc
138	19	1.5	88607	12	ADQ97648	Adq97648 Mouse can	18	1.4	555	10	ADH45899	Adh45899 Human lun
139	19	1.5	110000	2	AAT42063_13	Continuation (14 o	18	1.4	555	12	ADE71651	Ad71651 Human lun
140	19	1.5	110000	2	AAT42063_14	Continuation (15 o	18	1.4	555	13	ADJ19818	Adj19818 Human lun
141	19	1.5	110000	6	ABQ69245_14	Continuation (3 of	18	1.4	565	6	ABQ42635	Abq42635 Oligonuc1
142	19	1.5	110000	6	ABQ67195_2	Continuation (15 o	18	1.4	565	6	ABQ42634	Abq42634 Oligonuc1
143	19	1.5	139904	6	ABK83562	Abk83562 Human cDN	18	1.4	595	3	AAf11719	Aaf11719 Aspergill
144	19	1.5	157875	6	ABK99972	Abk99972 Human CAD	18	1.4	623	6	ABN61328	Abn61328 Human can
145	19	1.5	190000	10	ADL13752	Adl13752 Osteoarth	18	1.4	735	8	ABT20981	Abt20981 Aspergill
146	19	1.5	349980	13	ADT05648	Adt05648 Haemophil	18	1.4	735	8	ABT21175	Abt21175 Aspergill
147	19	1.5	349980	13	ADT05737	Adt05737 Haemophil	18	1.4	738	8	ABT19355	Abt19355 Aspergill
148	18	1.4	79	9	ADA60133	Ada60133 Soybean s	18	1.4	754	4	AA119169	Aa119169 Human bre
149	18	1.4	179	9	ADA60359	Ada60359 Soybean s	18	1.4	780	13	AAS10106	Ad10106 Human the
150	18	1.4	253	9	ADA60324	Ada60324 Soybean s	18	1.4	796	3	AA13767	Aa13767 Human ste
151	18	1.4	270	10	ABZ38894	Abz38894 N gonorr	18	1.4	796	3	AA13767	Aa13767 Human ste
152	18	1.4	276	10	ABZ38894	Abz38894 N gonorr	18	1.4	796	3	AA13767	Aa13767 Human ste
153	18	1.4	342	3	ABV07858	Abv07858 Human pro	18	1.4	799	4	AA13767	Aa13767 Human ste
154	18	1.4	345	4	AAK55429	Aak55429 Human imm	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
155	18	1.4	354	10	ACF67717	Acf67717 Phototrab	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
156	18	1.4	381	8	ABX52492	Abx52492 Soybean s	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
157	18	1.4	403	9	ADA60433	Ada60433 Soybean s	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
158	18	1.4	406	6	ABN64918	Abn64918 Human can	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
159	18	1.4	412	5	ABV61290	Abv61290 Human pro	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
160	18	1.4	432	3	AAAC36964	Aac36964 Arabidops	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
161	18	1.4	435	3	AAAC36964	Aac36964 Arabidops	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
162	18	1.4	451	3	AAAC27380	Aac27380 Human sec	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
163	18	1.4	495	2	AAV07332	Aav07332 Stem cell	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
164	18	1.4	495	2	AAV07332	Aav07332 Stem cell	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
165	18	1.4	498	2	AAV07332	Aav07332 Stem cell	18	1.4	820	2	AAQ11541	Aaq11541 Human ste
166	18	1.4	498	2	AAV07332	Aav07332 Stem cell	18	1.4	820	2	AAQ11541	Aaq11541 Human ste

C 240	18	1.4	820	12	ADP99315	Adp99315 Human ste	313	18	1.4	2834	3	AAC59739	Aac59739 Human sec
C 241	18	1.4	822	6	ABV78250	Abv78250 Human ste	314	18	1.4	2951	8	ABT18167	Abt18167 Aspergill
C 242	18	1.4	822	6	ABZ35826	Abz35826 Human ste	c 315	18	1.4	2969	8	ABX10380	Abx10380 DNA encod
C 243	18	1.4	822	6	ABX10069	Abx10069 Human ste	316	18	1.4	2979	8	ABT19793	Abt19793 Aspergill
C 244	18	1.4	822	6	ABL191791	Ab191791 Human pol	317	18	1.4	2996	8	ABT19787	Abt19787 Aspergill
C 245	18	1.4	828	6	ABQ19340	Abq19340 Oligonucle	318	18	1.4	2996	6	ABT19981	Abt19981 Aspergill
C 246	18	1.4	828	6	ABQ19341	Abq19341 Oligonucle	319	18	1.4	3052	6	ABK83776	Abk83776 Human CDN
C 247	18	1.4	860	8	ADA69923	Ada69923 Rice gene	c 320	18	1.4	3243	10	ADB63141	Adb63141 Human CDN
C 248	18	1.4	863	11	ACN82158	Acn82158 Breast ca	c 321	18	1.4	3303	6	AAS61867	Aas61867 Lung smal
C 249	18	1.4	864	2	AAQ20845	Aaq20845 Encodes h	c 322	18	1.4	3304	6	ABK83988	Abk83988 Human CDN
C 250	18	1.4	889	9	ACC84605	Acc84605 Human gln	c 323	18	1.4	3304	12	ADP43839	Adp43839 Human DNA
C 251	18	1.4	906	2	AAQ42808	Aaq42808 Human IL-	c 324	18	1.4	3310	6	ABK76552	Abk76552 cDNA enco
C 252	18	1.4	912	2	AAQ42809	Aaq42809 Human MGF	c 325	18	1.4	3310	13	ACN37823	Acn37823 Tumour-as
C 253	18	1.4	935	11	ACN90584	Acn90584 Breast ca	c 326	18	1.4	3310	13	ADR97393	Adr97393 Human XFC
C 254	18	1.4	942	9	ADA32208	Ada32208 DNA encod	c 327	18	1.4	3357	5	AAS72502	Aas72502 DNA encod
C 255	18	1.4	945	6	ABQ20416	Abq20416 Oligonucle	328	18	1.4	3402	2	AAH54452	Aah54452 S. epider
C 256	18	1.4	945	6	ABQ20417	Abq20417 Oligonucle	c 329	18	1.4	3462	2	AZ77533	Aaz77533 Human ova
C 257	18	1.4	951	8	ABT18761	Abt18761 Aspergill	c 330	18	1.4	3614	11	ADN95698	Adn95698 Human BSC
C 258	18	1.4	979	8	ABT18567	Abt18567 Aspergill	331	18	1.4	3668	4	AAH54402	Aah54402 S. epider
C 259	18	1.4	996	8	ABT20577	Abt20577 Aspergill	c 332	18	1.4	3806	6	ABT73857	Abt73857 Human SCF
C 260	18	1.4	996	8	ABT20383	Abt20383 Aspergill	c 333	18	1.4	3807	4	AAH41340	Aah41340 Human ste
C 261	18	1.4	1024	10	ABZ83697	Abz83697 Toxicolog	c 334	18	1.4	3807	4	AAH41340	Aah41340 Human ste
C 262	18	1.4	1059	2	AAT63113	Aat63113 Human sol	c 335	18	1.4	3807	4	AAH41340	Aah41340 Human ste
C 263	18	1.4	1078	12	ADO62584	Ado62584 Transcrip	c 336	18	1.4	3807	5	AAH41340	Aah41340 Human ste
C 264	18	1.4	1088	2	AAQ11543	Aaq11543 Human Ste	c 337	18	1.4	3807	5	AAH41340	Aah41340 Human ste
C 265	18	1.4	1088	2	AAT04891	Aat04891 Human SCF	c 338	18	1.4	3807	5	AAH41340	Aah41340 Human ste
C 266	18	1.4	1088	3	AAH13715	Aah13715 Human SCF	c 339	18	1.4	3807	5	AAH41340	Aah41340 Human ste
C 267	18	1.4	1088	4	AAH41345	Aah41345 Human SCF	c 340	18	1.4	3807	10	ADP52471	Adp52471 Human ste
C 268	18	1.4	1088	4	AAH41345	Aah41345 Human SCF	c 341	18	1.4	3807	12	ADP99313	Adp99313 Human ste
C 269	18	1.4	1088	4	AAH41345	Aah41345 Human SCF	c 342	18	1.4	4063	2	AAQ11540	Aaq11540 Human Ste
C 270	18	1.4	1088	5	AAH41345	Aah41345 Human SCF	c 343	18	1.4	4100	4	AAH41340	Aah41340 Human ste
C 271	18	1.4	1088	5	AAH41345	Aah41345 Human SCF	c 344	18	1.4	4335	8	ACA36923	Acc36923 Prokaryot
C 272	18	1.4	1088	5	AAH41345	Aah41345 Human SCF	c 345	18	1.4	4699	6	AAH41340	Aah41340 Human ste
C 273	18	1.4	1088	6	AAH41345	Aah41345 Human SCF	c 346	18	1.4	4736	3	AAH41340	Aah41340 Human ste
C 274	18	1.4	1088	6	ABT35478	Abt35478 Human SCF	347	18	1.4	5204	6	ABL32900	Ab132900 Human imm
C 275	18	1.4	1088	10	ABT35478	Abt35478 Human SCF	c 348	18	1.4	5263	4	ABL32900	Ab132900 Human imm
C 276	18	1.4	1088	12	ADP52490	Adp52490 Human ste	c 349	18	1.4	5416	4	ABL32900	Ab132900 Human imm
C 277	18	1.4	1088	12	ADP99332	Adp99332 Human ste	350	18	1.4	5447	4	AAH41340	Aah41340 Human ste
C 278	18	1.4	1089	2	AAT63114	Aat63114 Human sol	c 351	18	1.4	5571	6	ABQ70974	Abq70974 Listeria
C 279	18	1.4	1155	5	AAH72499	Aah72499 DNA encod	c 352	18	1.4	5864	4	AAH41340	Aah41340 Human ste
C 280	18	1.4	1215	8	ACA25257	Acc25257 Prokaryot	c 353	18	1.4	5864	4	AAH41340	Aah41340 Human ste
C 281	18	1.4	1227	6	ABN66135	Abn66135 Streptoco	c 354	18	1.4	5864	4	AAH41340	Aah41340 Human ste
C 282	18	1.4	1227	13	ADS58417	Ads58417 Bacterial	c 355	18	1.4	5864	4	AAH41340	Aah41340 Human ste
C 283	18	1.4	1256	9	ADA50160	Ada50160 Human CAR	c 356	18	1.4	5864	5	AAH41340	Aah41340 Human ste
C 284	18	1.4	1272	2	AAT63110	Aat63110 Human sol	c 357	18	1.4	5864	5	AAH41340	Aah41340 Human ste
C 285	18	1.4	1321	13	ADS88050	Ads88050 Tumour tr	c 358	18	1.4	5864	6	AAH41340	Aah41340 Human ste
C 286	18	1.4	1383	12	ADN74714	Adn74714 Thale cre	c 359	18	1.4	5864	10	ADE52475	Ades52475 Human ste
C 287	18	1.4	1404	2	AAQ11542	Aaq11542 Human Ste	c 360	18	1.4	5865	12	ADP99317	Adp99317 Human ste
C 288	18	1.4	1404	2	AAT04890	Aat04890 Human SCF	361	18	1.4	5883	6	ABL33732	Ab133732 Human imm
C 289	18	1.4	1404	3	AAH13714	Aah13714 Human SCF	c 362	18	1.4	5903	4	ABL20224	Ab120224 Drosophil
C 290	18	1.4	1404	4	AAH41344	Aah41344 Human SCF	363	18	1.4	5976	6	AAH41340	Aah41340 Human ste
C 291	18	1.4	1404	4	AAH41344	Aah41344 Human SCF	364	18	1.4	5976	6	ABK34007	Abk34007 Human DNA
C 292	18	1.4	1404	4	AAH41344	Aah41344 Human SCF	365	18	1.4	5976	10	ABE54106	Abes54106 Pretreate
C 293	18	1.4	1404	5	AAH41344	Aah41344 Human SCF	366	18	1.4	5976	10	ABE54234	Abes54234 Pretreate
C 294	18	1.4	1404	5	AAH41344	Aah41344 Human SCF	367	18	1.4	5976	10	ADE84096	Ades84096 Human lym
C 295	18	1.4	1404	5	AAH41344	Aah41344 Human SCF	368	18	1.4	5976	10	ADE84172	Ades84172 Human lym
C 296	18	1.4	1404	6	AAH41344	Aah41344 Human SCF	369	18	1.4	5976	13	ADS89538	Ads89538 Oligonucle
C 297	18	1.4	1404	6	AAH41344	Aah41344 Human SCF	370	18	1.4	5976	13	ADS89538	Ads89538 Oligonucle
C 298	18	1.4	1404	10	ADE52488	Ades52488 Human ste	371	18	1.4	6049	6	ABL32229	Ab132229 Human imm
C 299	18	1.4	1404	12	ADP99330	Adp99330 Human ste	c 372	18	1.4	6078	4	ABL33732	Ab133732 Human imm
C 300	18	1.4	1405	10	ACA56718	Acc56718 Human alg	373	18	1.4	6222	6	ABL32693	Ab132693 Human imm
C 301	18	1.4	1405	11	ADN95541	Adn95541 Human BSC	374	18	1.4	6255	6	ABL32960	Ab132960 Human imm
C 302	18	1.4	1405	12	ADI56514	Adi56514 Human pol	375	18	1.4	6351	6	ABL33174	Ab133174 Human imm
C 303	18	1.4	1461	4	AAE32431	Aae32431 LTM18 (li	376	18	1.4	6365	6	ABL32124	Ab132124 Human imm
C 304	18	1.4	1459	2	AAQ74818	Aaq74818 Firefly l	377	18	1.4	6487	9	ABL62896	Ab162896 Human alp
C 305	18	1.4	1735	6	ABE52000	Abes52000 cDNA enco	378	18	1.4	6533	6	ABK28384	Abk28384 DNA trans
C 306	18	1.4	1832	6	ABE55229	Abes55229 Pax prote	379	18	1.4	7028	4	AAH41340	Aah41340 Human ste
C 307	18	1.4	2000	8	ADA71913	Ada71913 Rice gene	380	18	1.4	7625	6	ABL33430	Ab133430 Human imm
C 308	18	1.4	2063	2	AAQ74819	Aaq74819 Firefly l	c 381	18	1.4	7703	4	AAH41340	Aah41340 Human ste
C 309	18	1.4	2166	4	AAH41345	Aah41345 Human SCF	382	18	1.4	7703	4	AAH41340	Aah41340 Human ste
C 310	18	1.4	2206	12	ADP04554	Adp04554 Sea squir	c 383	18	1.4	7703	5	AAH41340	Aah41340 Human ste
C 311	18	1.4	2616	8	ACA30186	Acc30186 Prokaryot	c 384	18	1.4	7703	5	AAH41340	Aah41340 Human ste
C 312	18	1.4	2793	4	AAH41345	Aah41345 Human SCF	c 385	18	1.4	7703	6	ABN90335	Abn90335 Human liv

386	18	1.4	7703	6	ABN90294	Human liv	459	18	1.4	110000	10	ACF65385_1	Continuation (2 of
387	18	1.4	7703	11	ADU15248	Human liv	c 460	18	1.4	110000	10	ACF65383_2	Continuation (3 of
388	18	1.4	7703	11	ADU15207	Human liv	c 461	18	1.4	110000	10	ACF67367_03	Continuation (4 of
389	18	1.4	7759	4	AAK77916	Human imm	462	18	1.4	110000	10	ACF67367_30	Continuation (13 o
390	18	1.4	7781	6	ABL33168	Human imm	463	18	1.4	110000	12	ADH77486_12	Continuation (31 o
391	18	1.4	7857	6	ABQ67075	Human ang	464	18	1.4	110000	12	ADJ25985_15	Continuation (16 o
392	18	1.4	8794	4	AAAS0465	DNA encod	465	18	1.4	110000	12	ADN97989_15	Continuation (16 o
393	18	1.4	8794	4	AAAL06245	Human rep	466	18	1.4	110000	12	ADN97989_15	Continuation (16 o
394	18	1.4	8962	6	ABL32686	Human imm	c 467	18	1.4	110000	13	ABD32629_1	Continuation (2 of
395	18	1.4	9741	6	ABL333323	Human imm	c 468	18	1.4	110000	13	ABD32629_2	Continuation (3 of
396	18	1.4	10855	6	ABL334388	Human imm	469	18	1.4	122614	11	ACN44098	Human gen
397	18	1.4	10886	6	ABL341134	Human imm	c 470	18	1.4	122923	11	ACN44026	Human gen
398	18	1.4	11155	6	ABL341134	Human imm	471	18	1.4	143391	10	ADL13648	Adl13648 Osteoarth
399	18	1.4	11754	6	AAAS04033	Human ABC	c 472	18	1.4	149034	11	ADP65433	Human ABC
400	18	1.4	11754	6	ABAL18638	Human ner	c 473	18	1.4	168575	4	AAH21613	AAH21613 Human hyp
401	18	1.4	13079	5	ABAL18638	Human ner	c 474	18	1.4	183999	4	AAQ92831	AAQ92831 Human ABC
402	18	1.4	13549	4	ABL07964	Human imm	475	18	1.4	185695	8	AAQ77405	AAQ77405 Human THB
403	18	1.4	14540	4	ABL07964	Human imm	c 476	18	1.4	185695	8	AAQ77405	AAQ77405 Human THB
404	18	1.4	14987	6	ABL32630	Human imm	477	18	1.4	201766	10	ADL13771	Adl13771 Osteoarth
405	18	1.4	15121	6	ABL32630	Human che	478	18	1.4	243390	13	ABD33366	Human car
406	18	1.4	15387	6	ABL32185	Human imm	c 479	18	1.4	285118	5	AAH41227	AAH41227 Pyrococcu
407	18	1.4	15789	4	AAAS33399	DNA encod	c 480	18	1.4	321491	11	ACN44202	ACN44202 Human gen
408	18	1.4	15789	4	AAAS35072	DNA #22 e	c 481	17	1.3	17	8	ABT37711	ABT37711 Tumour su
409	18	1.4	15789	10	ADL46514	Human imm	c 482	17	1.3	20	12	ADQ81511	ADQ81511 PCR prime
410	18	1.4	17213	6	ABL334482	Human imm	483	17	1.3	25	2	AAQ32035	AAQ32035 Oligonuc
411	18	1.4	17213	6	ABL334482	Human imm	c 484	17	1.3	25	2	AAQ28884	AAQ28884 SCF speci
412	18	1.4	17596	6	ABL32575	Human imm	c 485	17	1.3	51	4	AAH39640	AAH39640 Human SNP
413	18	1.4	17959	6	ABL32575	Human imm	c 486	17	1.3	141	13	ADS01678	Ads01678 Staphyloc
414	18	1.4	18283	6	ABL54342	Chemical	c 487	17	1.3	153	12	ADL87810	Adl87810 DNA up-re
415	18	1.4	18283	6	ABL70502	Chemical	c 488	17	1.3	153	12	ADL87809	Adl87809 DNA up-re
416	18	1.4	18283	6	ABL70502	Chemical	c 489	17	1.3	165	8	ABX35880	Abx35880 Bovine ES
417	18	1.4	18663	4	AAAK75859	Human imm	c 490	17	1.3	179	6	ABK29611	ABK29611 Colon ade
418	18	1.4	1993										







824	17	1.3	2648	10	ACA92448	Ac92448 DNA encod	897	17	1.3	4105	13	ADR87642	Adt87642 Human End
825	17	1.3	2660	6	ABQ34750	Human ova	898	17	1.3	4141	10	ADE57243	Ades7243 Rat gene
826	17	1.3	2708	10	ADQ90222	Novel hum	899	17	1.3	4179	4	ABL25562	Ab125562 Drosophil
827	17	1.3	2708	10	ADG90041	Human cdn	900	17	1.3	4182	6	AAD37230	Aad37230 Human dys
828	17	1.3	2727	4	AAI58340	Human pol	901	17	1.3	4202	12	ADQ22606	Adq22606 Human sof
829	17	1.3	2727	5	ADQ98548	DNA encod	902	17	1.3	4230	6	ABQ70965	Abq70965 Listeria
830	17	1.3	2727	9	ADQ98548	Human hum	903	17	1.3	4310	3	AAC77525	Aac77525 Human ORF
831	17	1.3	2760	6	ABN66563	Streptoco	904	17	1.3	4313	5	AAS74873	Aas74873 DNA encod
832	17	1.3	2828	4	AAK94249	Human ful	905	17	1.3	4385	8	ABQ83347	Abq83347 Human KLK
833	17	1.3	2828	12	ADL30823	Full leng	906	17	1.3	4415	6	ABL33907	Ab133907 Human imm
834	17	1.3	2844	3	ABX05212	Human nov	907	17	1.3	4415	6	ABL34603	Ab134603 Human met
835	17	1.3	2865	3	AAZ28209	Non-malto	908	17	1.3	4415	6	ABL70424	Ab170424 Chemicall
836	17	1.3	2898	6	ABK87938	Human arg	909	17	1.3	4415	7	ADS99864	Ads99864 Complemen
837	17	1.3	2898	12	ADU58907	Human arg	910	17	1.3	4415	10	ADF76792	Adf76792 Novel hum
838	17	1.3	2921	4	AAI60126	Human pol	911	17	1.3	4415	12	ADN05936	Adn05936 Antipsori
839	17	1.3	2930	4	ABL20678	Drosophil	912	17	1.3	4415	12	ADO20315	Ado20315 Human PRO
840	17	1.3	2933	5	AAE74797	DNA encod	913	17	1.3	4415	12	ADO20317	Ado20317 Human PRO
841	17	1.3	2940	11	ADN95667	Human BEC	914	17	1.3	4415	13	ADP55594	Adp55594 Human PRO
842	17	1.3	2941	10	ADU71105	Human met	915	17	1.3	4498	6	AAD37258	Aad37258 Adeno-ass
843	17	1.3	2952	10	ADA54019	Human cod	916	17	1.3	4617	12	ADQ20862	Adq20862 Human sof
844	17	1.3	2954	8	ABZ23169	Polynucle	917	17	1.3	4618	4	ABK42604	Abk42604 Genomic s
845	17	1.3	2954	12	ADQ83535	Human tum	918	17	1.3	4618	9	ADB60760	Adb60760 Connectiv
846	17	1.3	2954	13	ADQ84715	Human tum	919	17	1.3	4705	10	ADF74747	Adf74747 Murine dn
847	17	1.3	2954	13	ACN40459	Tumour-as	920	17	1.3	4707	12	ADQ29629	Adq29629 Human col
848	17	1.3	3025	4	AAI65200	Human neu	921	17	1.3	4718	4	AAH98223	Aah98223 Human EST
849	17	1.3	3031	3	AAAG64722	C. tracho	922	17	1.3	4718	10	ADI21512	Adi21512 Novel hum
850	17	1.3	3031	3	AAH56225	Chlamydia	923	17	1.3	4740	3	AAA95896	Aaa95896 Human KLK
851	17	1.3	3031	6	ABL92454	Chlamydia	924	17	1.3	4765	12	ADJ64312	Adj64312 Cartilage
852	17	1.3	3037	3	AAE74986	Human ORF	925	17	1.3	4778	2	AAI96562	Aai96562 Human neu
853	17	1.3	3081	6	ABQ74346	Human PNI	926	17	1.3	4825	6	AAD37257	Aad37257 Adeno-ass
854	17	1.3	3088	2	AAQ13727	Drosophil	927	17	1.3	4839	3	AAC47635	Aac47635 Arabidops
855	17	1.3	3096	4	ABL10095	Protein	928	17	1.3	4848	6	AAD37263	Aad37263 Adeno-ass
856	17	1.3	3100	10	AAI51702	Human nuc	929	17	1.3	4869	4	AAH18396	Aah18396 Human cdn
857	17	1.3	3104	3	AAA28215	Non-malto	930	17	1.3	4869	5	AAI93855	Aai93855 Human sto
858	17	1.3	3128	8	ABZ10017	Haematopo	931	17	1.3	4877	13	ADI61640	Adi61640 Human cdn
859	17	1.3	3128	8	ABZ10163	Haematopo	932	17	1.3	4898	12	ADQ25276	Adq25276 Human sof
860	17	1.3	3136	3	AAAO9386	Human DNA	933	17	1.3	4932	6	ABK46137	Abk46137 cDNA enco
861	17	1.3	3136	4	AAO98154	Mammalian	934	17	1.3	4933	8	AAI57279	Aai57279 KIAA0644
862	17	1.3	3143	4	ABL22778	Drosophil	935	17	1.3	4933	10	ACC72792	Acc72792 Human can
863	17	1.3	3215	11	ACN88795	Breast ca	936	17	1.3	4933	11	ADN95166	Adn95166 Human BPC
864	17	1.3	3224	3	AAA75848	Sequence	937	17	1.3	4933	11	ADN95093	Adn95093 Human LFC
865	17	1.3	3303	6	ABZ22372	Signal pa	938	17	1.3	4933	12	ADQ21195	Adq21195 Human sof
866	17	1.3	3361	10	ADU62810	Human cdn	939	17	1.3	4933	12	ADQ21195	Adq21195 Human sof
867	17	1.3	3369	1	AAAN81690	Atrial na	940	17	1.3	4941	13	ADP25678	Adp25678 Breast ca
868	17	1.3	3388	13	ADRO6827	Full leng	941	17	1.3	4941	12	ADPO9634	Adpo9634 Rice cinn
869	17	1.3	3391	4	ABL01932	Drosophil	942	17	1.3	4949	12	ADQ25090	Adq25090 Human sof
870	17	1.3	3531	6	ABD37238	Human dys	943	17	1.3	4966	6	AAD37256	Aad37256 Adeno-ass
871	17	1.3	3531	8	ABX63123	Human cdn	944	17	1.3	4990	6	AAD37262	Aad37262 Adeno-ass
872	17	1.3	3737	4	ABL11603	Drosophil	945	17	1.3	5060	6	AAD37264	Aad37264 Adeno-ass
873	17	1.3	3755	12	ADJ76120	Marker ge	946	17	1.3	5079	6	AAS94937	Aas94937 Human DNA
874	17	1.3	3778	13	ADRO8277	Full leng	947	17	1.3	5127	4	AAAS45425	Aas45425 Chemicall
875	17	1.3	3778	13	ADRO8277	Full leng	948	17	1.3	5149	6	AAD37255	Aad37255 Adeno-ass
876	17	1.3	3812	2	AAI42219	Full leng	949	17	1.3	5174	13	ADT05447	Adt05447 Haemophil
877	17	1.3	3812	2	AAI79605	TATA-bind	950	17	1.3	5244	6	ABK40013	Abk40013 Human che
878	17	1.3	3825	6	ABK72928	Bacillus	951	17	1.3	5276	6	ABU32151	Abu32151 Human imm
879	17	1.3	3837	12	ADJ12448	DNA fragm	952	17	1.3	5276	12	ADN05527	Adn05527 Antipsori
880	17	1.3	3857	4	AAH54116	S. epide	953	17	1.3	5287	6	ABL33493	Ab133493 Human imm
881	17	1.3	3858	6	AAAD37237	Human dys	954	17	1.3	5303	6	ABL32870	Ab132870 Human imm
882	17	1.3	3941	8	ADAB3888	Human MUC	955	17	1.3	5303	6	ABL33339	Ab133339 Brassica
883	17	1.3	3941	13	ADQ80265	Mucin 3B	956	17	1.3	5313	6	ABL32763	Ab132763 Human imm
884	17	1.3	3972	12	ADQ63887	Novel hum	957	17	1.3	5317	6	ABL32609	Ab132609 Human imm
885	17	1.3	3999	6	AAD37234	Human dys	958	17	1.3	5339	6	ABK81998	Abk81998 DNA encod
886	17	1.3	4011	12	ADQ03170	Corn orth	959	17	1.3	5365	6	ABL32047	Ab132047 Human imm
887	17	1.3	4079	10	ACA56659	Human eig	960	17	1.3	5365	6	ABL70118	Ab170118 Chemicall
888	17	1.3	4079	12	ADI56455	Human pol	961	17	1.3	5365	6	AAE61050	Aae61050 Human gen
889	17	1.3	4097	8	ABX05211	Human nov	962	17	1.3	5366	4	ABL03676	Ab103676 Drosophil
890	17	1.3	4105	3	AAQ34583	ETA recep	963	17	1.3	5439	6	ABL33403	Ab133403 Human imm
891	17	1.3	4105	3	AAA38341	Human end	964	17	1.3	5462	6	ABL33916	Ab133916 Human imm
892	17	1.3	4105	6	ABZ35259	Human gen	965	17	1.3	5464	6	ABK31452	Abk31452 Signal tr
893	17	1.3	4105	6	ABV94238	Breast ca	966	17	1.3	5464	6	ABL70429	Ab170429 Chemicall
894	17	1.3	4105	8	ABZ42662	Human end	967	17	1.3	5471	4	ABL10094	Ab110094 Drosophil
895	17	1.3	4105	12	ADF42785	Human end	968	17	1.3	5551	6	ABL70157	Ab170157 Chemicall
896	17	1.3	4105	12	ADQ18007	Human sof	969	17	1.3	5551	6	AAE61099	Aae61099 Human gen

970 17 1.3 5552 6 ABL33258 Human imm  
971 17 1.3 5573 12 Adq64971 Novel hum  
972 17 1.3 5640 10 ADB84161 Gene expr  
973 17 1.3 5659 6 ABL32374 Human imm  
974 17 1.3 5659 6 ABK33931 Human DNA  
975 17 1.3 5659 6 ABL34486 Human met  
976 17 1.3 5659 7 Ad99747 Bisulphit  
977 17 1.3 5659 8 ADA20366 Prostata  
978 17 1.3 5659 8 ADA84173 Human ren  
979 17 1.3 5735 6 ABK94926 Human nov  
980 17 1.3 5735 6 ABK94980 Human nov  
981 17 1.3 5759 6 ABL33704 Human imm  
982 17 1.3 5768 6 ABL34162 Human imm  
983 17 1.3 5771 6 ABL33155 Human che  
984 17 1.3 5807 6 ABL33155 Human che  
985 17 1.3 5815 6 ABK40023 DNA trans  
986 17 1.3 5823 6 ABL28382 Human imm  
987 17 1.3 5890 6 ABL33338 Human imm  
988 17 1.3 5952 5 AAD06794 Human dys  
989 17 1.3 5971 6 ABL32383 Human imm  
990 17 1.3 6014 6 ABL32383 Human imm  
991 17 1.3 6014 6 ABL32383 Human imm  
992 17 1.3 6014 6 ABL32383 Human imm  
993 17 1.3 6024 6 ABL33646 Human imm  
994 17 1.3 6025 4 ABL29649 Drosophil  
995 17 1.3 6028 10 ADE84205 Human lym  
996 17 1.3 6028 10 ADE84129 Human lym  
997 17 1.3 6042 6 ABL33944 Human imm  
998 17 1.3 6044 6 ABL11602 Drosophil  
999 17 1.3 6049 6 ABL32373 Human imm  
1000 17 1.3 6061 4 AAS45335 Chemical

ALIGNMENTS

RESULT 1  
AAC67281  
ID AAC67281 standard; cDNA; 1274 BP.  
XX AC AAC67281;  
XX AC AAC67281;  
XX 09-APR-2001 (first entry)  
XX Human Siah-1alpha coding sequence SEQ ID NO: 1.  
XX Human; protein degradation; siah-mediated degradation protein; SMDP;  
XX SCF-complex protein; SCP; siah-1 interacting protein; SIP;  
XX Skpl-associated F-box protein; SAF-1; SAF-2; SAD; cancer; cell division;  
XX Skpl-associated destruction-box protein; inflammatory disease; ss.  
XX Homo sapiens.  
XX WO200077207-A2.  
XX 21-DEC-2000.  
XX 09-JUN-2000; 2000WO-US015873.  
XX 11-JUN-1999; 99US-00330517.  
XX (BURN-) BURNHAM INST.  
XX Reed JC, Matsuzawa S;  
XX WPI; 2001-071273/08.  
XX P-PSDB; AAB35157.  
XX Siah-Mediated Degradation Protein, useful for drug screening, for  
XX therapeutic applications and for functional genomics.  
XX Claim 5; Page 95-97; 121pp; English.

CC The present invention provides the protein and coding sequences of  
CC several siah-mediated degradation proteins and SCF-complex proteins.  
CC These are designated siah-1alpha, Siah-1 interacting protein (SIP), which  
CC encodes two proteins due to alternative splicing (SIP-L and SIP-S), Skpl-  
CC associated F-box protein-lalpha and beta and -2 (SAF-lalpha, SAF-beta  
CC and SAF-2) and Skpl-associated destruction-box protein (SAD). The  
CC proteins and their coding sequences are useful in the diagnosis and  
CC treatment of cancers, disorders where too little cell division occurs  
CC such as bone marrow aplasia, immunodeficiencies and inflammatory  
CC diseases including sepsis, fibrosis, arthritis and graft versus host  
CC disease  
XX SQ Sequence 1274 BP; 336 A; 273 C; 253 G; 412 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1274; DB 4; Length 1274;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTAGTTGTTTATGGTCCATTTCTATTTAGCATTTATTTCTATGTAGTCTAT 60  
DB 1 TTTCTTAGTTGTTTATGGTCCATTTCTATTTAGCATTTATTTCTATGTAGTCTAT 60  
QY 61 CCAAGACGATTAAGGGAGTTCCACATGTTTCCGGAACATTTGAAAAGAGAGCTTATC 120  
DB 61 CCAAGACGATTAAGGGAGTTCCACATGTTTCCGGAACATTTGAAAAGAGAGCTTATC 120  
QY 121 CAGTGTACAGATCCTAATAAGTCACATTCAGTGTAAATTTTATTTTATTTTATTTT 180  
DB 121 CAGTGTACAGATCCTAATAAGTCACATTCAGTGTAAATTTTATTTTATTTTATTTT 180  
QY 181 TTAATCTTATTTTCTTCT 240  
DB 181 TTAATCTTATTTTCTTCT 240  
QY 241 TATGCGATGTAACATTTATTTAAGTAAGTCAATGTTTATTTTATTTTCTCTCTCTCT 300  
DB 241 TATGCGATGTAACATTTATTTAAGTAAGTCAATGTTTATTTTATTTTCTCTCTCTCT 300  
QY 301 CTTATGTATTTTATTTTACAAATGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360  
DB 301 CTTATGTATTTTATTTTACAAATGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360  
QY 361 TCGAAGTGTCCACATCCAGAGGGT 420  
DB 361 TCGAAGTGTCCACATCCAGAGGGT 420  
QY 421 GACTTGGCGAGTCTTTTGT 480  
DB 421 GACTTGGCGAGTCTTTTGT 480  
QY 481 CAATGTACAGATGAGCGTCTTTGT 540  
DB 481 CAATGTACAGATGAGCGTCTTTGT 540  
QY 541 ACTTGGCGGGGCGCTTTTGGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAATTTCA 600  
DB 541 ACTTGGCGGGGCGCTTTTGGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAATTTCA 600  
QY 601 GTACTTTTCCCTGTAAATATATCGGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
DB 601 GTACTTTTCCCTGTAAATATATCGGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
QY 661 AAAGCAGACATGAAGAGCTCTGT 720  
DB 661 AAAGCAGACATGAAGAGCTCTGT 720  
QY 721 TCCGTGTAATGGCAAGGCTCTCTGT 780  
DB 721 TCCGTGTAATGGCAAGGCTCTCTGT 780  
QY 781 TCCATTACACCTTACAGGAGGATATAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840  
DB 781 TCCATTACACCTTACAGGAGGATATAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840

QY 841 GGTGCTGTTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTCACTTCATGTTAGTCTTA 900  
DB 841 GGTGCTGTTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTCACTTCATGTTAGTCTTA 900  
QY 901 GAGAAACAGGAAATACGATGTCACAGCAGTCTTCCGAATCGTACAGCTGATAGGA 960  
DB 901 GAGAAACAGGAAATACGATGTCACAGCAGTCTTCCGAATCGTACAGCTGATAGGA 960  
QY 961 ACACGCAAGCAAGCTGAAAATTTTGTCTTACCGACTTGAGCTAAATGGTCATAGCGCAGCA 1020  
DB 961 ACACGCAAGCAAGCTGAAAATTTTGTCTTACCGACTTGAGCTAAATGGTCATAGCGCAGCA 1020  
QY 1021 TTGACTTGGGAAGCGACTCTCGATCTATTCATGAAGGAATTGCAACAGCCATTATGAAT 1080  
DB 1021 TTGACTTGGGAAGCGACTCTCGATCTATTCATGAAGGAATTGCAACAGCCATTATGAAT 1080  
QY 1081 AGCGACTGTCAGTCTTTCACACACGATTTGACAGCTTTTTCGAGAAAATGCAATTTTA 1140  
DB 1081 AGCGACTGTCAGTCTTTCACACACGATTTGACAGCTTTTTCGAGAAAATGCAATTTTA 1140  
QY 1141 GGCATCAATGTAATTTCCATGTTGAAATGGCAATCAAACTTTTCTGCGCCAGTGT 1200  
DB 1141 GGCATCAATGTAATTTCCATGTTGAAATGGCAATCAAACTTTTCTGCGCCAGTGT 1200  
QY 1201 TTAAACTTCAGTTTCACAGAAAATAGGCAACCCATCTGTCTGCCAACCTTAACTCTTTT 1260  
DB 1201 TTAAACTTCAGTTTCACAGAAAATAGGCAACCCATCTGTCTGCCAACCTTAACTCTTTT 1260  
QY 1261 CGGTAGGTGGAAGC 1274  
DB 1261 CGGTAGGTGGAAGC 1274

## RESULT 2

ABV28804  
ID ABV28804 standard; cDNA; 2924 BP.

AC ABV28804;

XX 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 28795.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US045171.

XX 17-FEB-2000; 2000US-0183019P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207854P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219407P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JB;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 6058; 11750pp; English.

XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing whether prostate cancer has metastasized in a patient; (g) determining the aggressiveness or indolence of prostate cancer in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 2924 BP; 871 A; 562 C; 562 G; 925 T; 0 U; 4 Other;

Query Match 100.0%; Score 1274; DB 5; Length 2924;

Best Local Similarity 100.0%; Pred No. 0;

Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCTTTAGTTTATGCTCCATTTCTATTTTAGCATTTATTTCTATGTAGTCTAT 60

DB 480 TTTCTTTAGTTTATGCTCCATTTCTATTTTAGCATTTATTTCTATGTAGTCTAT 539

QY 61 CCAAGACGATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120

DB 540 CCAAGACGATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 599

QY 121 CAGTGTACAGATCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATCTTTT 180

DB 600 CAGTGTACAGATCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATCTTTT 659

QY 181 TTAATCTTATTTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAACTTTAAAGGACT 240

DB 660 TTAATCTTATTTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAACTTTAAAGGACT 719

QY 241 TATGCGATGTAACATTTATTAAGTAAGTCATGGTTATTAATTTTCTCTCTCTCT 300

DB 720 TATGCGATGTAACATTTATTAAGTAAGTCATGGTTATTAATTTTCTCTCTCTCT 779

QY 301 CCTATGATTTTATTTTCAAGATGAGCGCTCAGACTGCTACAGCATTTACTACCGGTACC 360

DB 780 CCTATGATTTTATTTTCAAGATGAGCGCTCAGACTGCTACAGCATTTACTACCGGTACC 839

QY 361 TCGAAGTGTCCACCATCCAGAGGTTGCTGCCCTGACTGGCACAACCTGCAATCAACAAT 420

DB 840 TCGAAGTGTCCACCATCCAGAGGTTGCTGCCCTGACTGGCACAACCTGCAATCAACAAT 899

QY 421 GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTTTGACTATGTGTACGCCCATCTTT 480

DB 900 GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTTTGACTATGTGTACGCCCATCTTT 959

QY 481 CAATGTCAGAGTGGCCATCTTGTGTTAGCAATTTGTCGCCCAAGCTCATGTTGTCCA 540

DB 960 CAATGTCAGAGTGGCCATCTTGTGTTAGCAATTTGTCGCCCAAGCTCATGTTGTCCA 1019

QY 541 ACTTGGCGGGGCCCTTTGGGATCCATTTCCCAACTTTGCTATGAGAGAAAGTGGCTAATCA 600

DB 1020 ACTTGGCGGGGCCCTTTGGGATCCATTTCCCAACTTTGCTATGAGAGAAAGTGGCTAATCA 1079

QY 601 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATATCTCTGCCACACACAGAA 660

DB 1080 GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATATCTCTGCCACACACAGAA 1139

QY 661 AAAGCAGACCATGAAGAGCTCTGAGTTTATGCGCTTATTTCTGCTCCGCTGGTGTCT 720

DB 1140 AAAGCAGACCATGAAGAGCTCTGAGTTTATGCGCTTATTTCTGCTCCGCTGGTGTCT 1199

QY 721 TCCCTGTAATGCAAGGCTCTCTGATGCTGTAATGCCCTCTCTGATGATCAGCATAAG 780

DB 1200 TCCCTGTAATGCAAGGCTCTCTGATGCTGTAATGCCCTCTCTGATGATCAGCATAAG 1259

QY 781 TCCATTACAACCCCTACAGGAGGATATAGTTTTTCTTGTGTACAGACATTAATCTTCCT 840

Db 1260 TCCATTACACCCCTACAGGAGAGATATAGTTTCTTGTACAGACATTAATCTTCCT 1319

Qy 841 GGTGCTGTTGACTGGGTGATGATGAGTCCCTGTTTGGCTTTCACTTCATGTTAGTCTTA 900

Db 1320 GGTGCTGTTGACTGGGTGATGATGAGTCCCTGTTTGGCTTTCACTTCATGTTAGTCTTA 1379

Qy 901 GAGAAACAGGAAATACAGTGTACAGGAGTCTTCCGCAATCGTACAGCTGATAGGA 960

Db 1380 GAGAAACAGGAAATACAGTGTACAGGAGTCTTCCGCAATCGTACAGCTGATAGGA 1439

Qy 961 ACACCCAGCAGCTGAAATTTTCTACCGACTTGGAGCTTAAATGGTCAATAGGCGACGA 1020

Db 1440 ACACCCAGCAGCTGAAATTTTCTACCGACTTGGAGCTTAAATGGTCAATAGGCGACGA 1499

Qy 1021 TTGACTTTGGGAAGCGACTCTCGATCTATTCATGAAGAAATGCAACAGCCATATGAAAT 1080

Db 1500 TTGACTTTGGGAAGCGACTCTCGATCTATTCATGAAGAAATGCAACAGCCATATGAAAT 1559

Qy 1081 AGCGACTGTCTAGTCTTTGACACCGACTTGCACAGCTTTTGGCAAAATGGCAATTTA 1140

Db 1560 AGCGACTGTCTAGTCTTTGACACCGACTTGCACAGCTTTTGGCAAAATGGCAATTTA 1619

Qy 1141 GGCATCAATGTAATATTTCCATGTTTGAATGGCAATCAAAATTTTCTGCCAGTGT 1200

Db 1620 GGCATCAATGTAATATTTCCATGTTTGAATGGCAATCAAAATTTTCTGCCAGTGT 1679

Qy 1201 TTAATACTTCAGTTTCCAGAAATTAAGGCAACCCATCTGTCTGCCAATCTAAATCTTT 1260

Db 1680 TTAATACTTCAGTTTCCAGAAATTAAGGCAACCCATCTGTCTGCCAATCTAAATCTTT 1739

Qy 1261 CGGTAGTGGAGC 1274

Db 1740 CGGTAGTGGAGC 1753

RESULT 3

ABV25468

ID ABV25468 standard; cDNA; 2924 BP.

XX AC ABV25468;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 25459.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PP 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0199862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 5054-5055; 11750pp; English.

PS The invention relates to an isolated nucleic acid molecule (I) comprising

XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy

CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 2924 BP; 871 A; 562 C; 562 G; 925 T; 0 U; 4 Other;

SQ

Query Match 100.0%; Score 1274; DB 5; Length 2924;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTCTTTAGTTGTTTATGGTCCATTTTCTATTTTAGCATTTATTTCTATGTAGTCTAT 60

Db 480 TTTCTTTAGTTGTTTATGGTCCATTTTCTATTTTAGCATTTATTTCTATGTAGTCTAT 539

Qy 61 CCAAAGCAGATTAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120

Db 540 CCAAAGCAGATTAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 599

Qy 121 CAGTGTACAGATCTTAATAAGTGTACATTCAGTGTAAATTTTATTTTAAATATCTTTT 180

Db 600 CAGTGTACAGATCTTAATAAGTGTACATTCAGTGTAAATTTTATTTTAAATATCTTTT 659

Qy 181 TTAATCTTATTTTCT 240

Db 660 TTAATCTTATTTTCT 719

Qy 241 TATGCGATGTAAACATTTATTAAGTGTATGTTTATTAATTTTCTCTCTCTCTCTCTCT 300

Db 720 TATGCGATGTAAACATTTATTAAGTGTATGTTTATTAATTTTCTCTCTCTCTCTCTCT 779

Qy 301 CCTTATGTTATTTTATTTTCAAGATGAGCGTGTACAGATTTACCTACCGGTACC 360

Db 780 CCTTATGTTATTTTATTTTCAAGATGAGCGTGTACAGATTTACCTACCGGTACC 839

Qy 361 TCGAAGTGTCCACCATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

Db 840 TCGAAGTGTCCACCATCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899

Qy 421 GACTTGGCGAGTCTTTTGGAGTGTCCAGTGTCTGTTGACTATGTTTACCGCCCATTCCT 480

Db 900 GACTTGGCGAGTCTTTTGGAGTGTCCAGTGTCTGTTGACTATGTTTACCGCCCATTCCT 959

Qy 481 CAATGTCCAGTGGCCATCTCTGTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCA 540

Db 960 CAATGTCCAGTGGCCATCTCTGTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCA 1019

Qy 541 ACTTGGCGGCGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAAGTGGCTAATTTCA 600

Db 1020 ACTTGGCGGCGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAAGTGGCTAATTTCA 1079

Qy 601 GTACTTTTCCCTGTAAATATGCGTCTTCTGATGTGAATTAATCTGCGCACACAGAA 660

Db 1080 GTACTTTTCCCTGTAAATATGCGTCTTCTGATGTGAATTAATCTGCGCACACAGAA 1139

Qy 661 AAAGCAGACATCAAGAGCTCTGTGAGTTTAGGCTTTATTCCTGTCGCTGCTGCTGCTGCT 720

Db 1140 AAAGCAGACATCAAGAGCTCTGTGAGTTTAGGCTTTATTCCTGTCGCTGCTGCTGCTGCT 1199

Qy 721 TCTGTAAATGGCAAGGCTCTCTGATGTCTGTAATGCCCTCCCTGATCATGACATAAG 780

Db 1200 TCTGTAAATGGCAAGGCTCTCTGATGTCTGTAATGCCCTCCCTGATCATGACATAAG 1259



QY 781 TCCATTACCAACCTACAGGAGGATATAGTTTTCTTGCTACAGACATTAATCTTCT 840  
DB 1260 TCCATTACCAACCTACAGGAGGATATAGTTTTCTTGCTACAGACATTAATCTTCT 1319  
QY 841 GGTGCTGTTGACTGGGTGATGATGCTGCTGTTTGGCTTTTCACTTCACTGTTAGTCTTA 900  
DB 1320 GGTGCTGTTGACTGGGTGATGATGCTGCTGTTTGGCTTTTCACTTCACTGTTAGTCTTA 1379  
QY 901 GAGAAACAGGAAATACAGATGTCACCGAGCTTCTTCCCAATCGTACAGCTGATGGA 960  
DB 1380 GAGAAACAGGAAATACAGATGTCACCGAGCTTCTTCCCAATCGTACAGCTGATGGA 1439  
QY 961 ACACGACAGCTGAAATTTTGTCTTACGACCTTGAAGCTAAATGGTCATAGGCGACGA 1020  
DB 1440 ACACGACAGCTGAAATTTTGTCTTACGACCTTGAAGCTAAATGGTCATAGGCGACGA 1499  
QY 1021 TTGACTTGGGAACGACTCCTCGATCTATTTCATGAAGGAATTGCAACAGCCATTATGAAT 1080  
DB 1500 TTGACTTGGGAACGACTCCTCGATCTATTTCATGAAGGAATTGCAACAGCCATTATGAAT 1559  
QY 1081 AGCGACTGTCTAGCTTTTGACACACGACATTTGACAGCTTTTTCGAGAAATGCAATTTA 1140  
DB 1560 AGCGACTGTCTAGCTTTTGACACACGACATTTGACAGCTTTTTCGAGAAATGCAATTTA 1619  
QY 1141 GGCATCAATGTAATATATCCATGTTGAAATGGCAATCAAAATTTTCTGCCAGTGT 1200  
DB 1620 GGCATCAATGTAATATATCCATGTTGAAATGGCAATCAAAATTTTCTGCCAGTGT 1679  
QY 1201 TTAATAACTTCAGTTTCACAGAAATAAGGACCCATCTGCTGCCAACCTTAAACTCTTT 1260  
DB 1680 TTAATAACTTCAGTTTCACAGAAATAAGGACCCATCTGCTGCCAACCTTAAACTCTTT 1739  
QY 1261 CGGTAGTGGAGC 1274  
DB 1740 CGGTAGTGGAGC 1753

## RESULT 4

ABV22970  
ID ABV22970 standard; cdna; 2924 BP.  
XX AC ABV22970;  
XX DT 13-SEP-2002 (first entry)  
XX DE Human prostate expression marker cdna 22961.  
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX KW pharmacogenomic marker; gene; ss.  
XX OS Homo sapiens.  
XX PN WO200160860-A2.  
XX PD 23-AUG-2001.  
XX PF 20-FEB-2001; 2001WO-US005171.  
XX PR 17-FEB-2000; 2000US-0183319P.  
XX PR 16-MAR-2000; 2000US-0189862P.  
XX PR 25-MAY-2000; 2000US-0207454P.  
XX PR 09-JUN-2000; 2000US-0211314P.  
XX PR 18-JUL-2000; 2000US-0219007P.  
XX PR 13-DEC-2000; 2000US-0255281P.  
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX PI Schlegel R, Endege WO, Monahan JB;  
XX DR WPI; 2001-662795/76.  
XX PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX Claim 1; Page 4080-4081; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX Sequence 2924 BP; 871 A; 562 C; 925 G; 925 T; 0 U; 4 Other;

Query Match 100.0%; Score 1274; DB 5; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTTAGTTGTTTATGCTCCATTTTCTATTTTAGCATTTATTTCTATGCTAT 60  
DB 480 TTTCTTTAGTTGTTTATGCTCCATTTTCTATTTTAGCATTTATTTCTATGCTAT 539  
QY 61 CCAAGACGATTAAGGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120  
DB 540 CCAAGACGATTAAGGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 599  
QY 121 CAGTGTACAGATCCTAAATAAGTGCACATTCAGTGTAAATTTATTTTAAATATCTTT 180  
DB 600 CAGTGTACAGATCCTAAATAAGTGCACATTCAGTGTAAATTTATTTTAAATATCTTT 659  
QY 181 TTAATCCTATTTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAACCTTAAAGGACT 240  
DB 660 TTAATCCTATTTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAACCTTAAAGGACT 719  
QY 241 TATGCATGTAACCATTTATTAAGTAAGTCATGTTATTAATTTTCTCTCTGCT 300  
DB 720 TATGCATGTAACCATTTATTAAGTAAGTCATGTTATTAATTTTCTCTCTGCT 779  
QY 301 CTTATGATTTATTTTTCAGAAATGAGCGGTACAGCTGCTACAGCATTTACCTACCGTACC 360  
DB 780 CTTATGATTTATTTTTCAGAAATGAGCGGTACAGCTGCTACAGCATTTACCTACCGTACC 839  
QY 361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCCCTGACGTCGGCAGCACTGCATCCACAAT 420  
DB 840 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCCCTGACGTCGGCAGCACTGCATCCACAAT 899  
QY 421 GACTTGGCGAGTCTTTTGGAGTCCAGTCTCTTTGACTATGTGTACCGCCATCTT 480  
DB 900 GACTTGGCGAGTCTTTTGGAGTCCAGTCTCTTTGACTATGTGTACCGCCATCTT 959  
QY 481 CAATGTCAGAGTGGCCATCTTGTGTTAGCAATCTGCGCCAAAGCTCATGTTGTCCA 540  
DB 960 CAATGTCAGAGTGGCCATCTTGTGTTAGCAATCTGCGCCAAAGCTCATGTTGTCCA 1019  
QY 541 ACTTCCCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATCGAGAAAGTGGCTAATTC 600  
DB 1020 ACTTCCCGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATCGAGAAAGTGGCTAATTC 1079  
QY 601 GTACTTTTCCCTGTAATATATGCGTCTTCTGATGTGAATTAATCTGCGCACACAGAA 660  
DB 1080 GTACTTTTCCCTGTAATATATGCGTCTTCTGATGTGAATTAATCTGCGCACACAGAA 1139  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCTTATTTCTGCGGCTGCTGCT 720  
DB 1140 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCTTATTTCTGCGGCTGCTGCT 1199  
QY 721 TCCTGTAATGCGCAAGGCTCTCTGGAATGCTGTAATGCTGTAATGCTGTAATGCTGTAATG 780



Db	1200	TCCTGTAATGGCAAGGCTCTCTGAGTGTGTAATGGCCCACTGATGCATCAGCATAG	1259	PR	07-JUL-2000;	2000US-0216880P.
				PR	11-JUL-2000;	2000US-0217487P.
				PR	11-JUL-2000;	2000US-0217496P.
QY	781	TCATTACAAACCTTACAGGAGAGATAGTTTTTCTTGCTACAGACATTAATCTTCCT	840	PR	11-JUL-2000;	2000US-0218290P.
				PR	26-JUL-2000;	2000US-0220963P.
Db	1260	TCATTACAAACCTTACAGGAGAGATAGTTTTTCTTGCTACAGACATTAATCTTCCT	1319	PR	26-JUL-2000;	2000US-0220964P.
				PR	14-AUG-2000;	2000US-0224518P.
QY	841	GGTGTGTGACTGGGTGATGATGACCTCTGTTTTGGCTTTCATCTCATGTTAGTCTTA	900	PR	14-AUG-2000;	2000US-0224519P.
				PR	14-AUG-2000;	2000US-0225213P.
Db	1320	GGTGTGTGACTGGGTGATGATGACCTCTGTTTTGGCTTTCATCTCATGTTAGTCTTA	1379	PR	14-AUG-2000;	2000US-0225214P.
				PR	14-AUG-2000;	2000US-0225266P.
QY	901	GAGAACAGGAAATACAGTGGTCCACGACAGTTCTTCGCAATCGTACAGCTGATAGGA	960	PR	14-AUG-2000;	2000US-0225267P.
				PR	14-AUG-2000;	2000US-0225268P.
Db	1380	GAGAACAGGAAATACAGTGGTCCACGACAGTTCTTCGCAATCGTACAGCTGATAGGA	1439	PR	14-AUG-2000;	2000US-0225270P.
				PR	14-AUG-2000;	2000US-0225447P.
QY	961	ACACGCAAGCAAGCTGAAATTTTGTCTACCGACTTGAGCTAAATGGTCATAGGCGACGA	1020	PR	14-AUG-2000;	2000US-0225757P.
				PR	14-AUG-2000;	2000US-0225758P.
Db	1440	ACACGCAAGCAAGCTGAAATTTTGTCTACCGACTTGAGCTAAATGGTCATAGGCGACGA	1499	PR	14-AUG-2000;	2000US-0225759P.
				PR	18-AUG-2000;	2000US-0226279P.
QY	1021	TTGACTTGGGAAGCACTCCTCGATCTATTCTATGAGCAATTTGCAACAGCCATTATGAAT	1080	PR	22-AUG-2000;	2000US-0226681P.
				PR	22-AUG-2000;	2000US-0226686P.
Db	1500	TTGACTTGGGAAGCACTCCTCGATCTATTCTATGAGCAATTTGCAACAGCCATTATGAAT	1559	PR	22-AUG-2000;	2000US-0227182P.
				PR	23-AUG-2000;	2000US-0227009P.
QY	1081	AGCGACTGTAGTCTTTGCACACCAAGCATTTGCACAGCTTTTTCGAGCAATATGCAATTTA	1140	PR	30-AUG-2000;	2000US-0228924P.
				PR	01-SEP-2000;	2000US-0229287P.
Db	1560	AGCGACTGTAGTCTTTGCACACCAAGCATTTGCACAGCTTTTTCGAGCAATATGCAATTTA	1619	PR	01-SEP-2000;	2000US-0229343P.
				PR	01-SEP-2000;	2000US-0229344P.
QY	1141	GGCATCAATGTAACTATTTCCATGTTGCAATGGCAATCAAACTTTCTGGCCAGTGT	1200	PR	01-SEP-2000;	2000US-0229345P.
				PR	05-SEP-2000;	2000US-0229509P.
Db	1620	GGCATCAATGTAACTATTTCCATGTTGCAATGGCAATCAAACTTTCTGGCCAGTGT	1679	PR	05-SEP-2000;	2000US-0229513P.
				PR	06-SEP-2000;	2000US-0230437P.
QY	1201	TTAAACTTCAGTTTCACAGAAATTAAGCACCCATCTGCTGCCAACCTTAAACTCTTT	1260	PR	06-SEP-2000;	2000US-0230438P.
				PR	08-SEP-2000;	2000US-0231242P.
Db	1680	TTAAACTTCAGTTTCACAGAAATTAAGCACCCATCTGCTGCCAACCTTAAACTCTTT	1739	PR	08-SEP-2000;	2000US-0231243P.
				PR	08-SEP-2000;	2000US-0231244P.
QY	1261	CGGTAGTGGAGAC 1274		PR	08-SEP-2000;	2000US-0231413P.
				PR	08-SEP-2000;	2000US-0231414P.
Db	1740	CGGTAGTGGAGAC 1753		PR	08-SEP-2000;	2000US-0232080P.
				PR	08-SEP-2000;	2000US-0232081P.
				PR	12-SEP-2000;	2000US-0231968P.
				PR	14-SEP-2000;	2000US-0232397P.
				PR	14-SEP-2000;	2000US-0232398P.
				PR	14-SEP-2000;	2000US-0232399P.
				PR	14-SEP-2000;	2000US-0232400P.
				PR	14-SEP-2000;	2000US-0232401P.
				PR	14-SEP-2000;	2000US-0233063P.
				PR	14-SEP-2000;	2000US-0233064P.
				PR	14-SEP-2000;	2000US-0233065P.
				PR	21-SEP-2000;	2000US-0234223P.
				PR	21-SEP-2000;	2000US-0234274P.
				PR	25-SEP-2000;	2000US-0234997P.
				PR	25-SEP-2000;	2000US-0234998P.
				PR	26-SEP-2000;	2000US-0234984P.
				PR	27-SEP-2000;	2000US-0235834P.
				PR	27-SEP-2000;	2000US-0235836P.
				PR	29-SEP-2000;	2000US-0236327P.
				PR	29-SEP-2000;	2000US-0236368P.
				PR	29-SEP-2000;	2000US-0236369P.
				PR	29-SEP-2000;	2000US-0236370P.
				PR	02-OCT-2000;	2000US-0236802P.
				PR	02-OCT-2000;	2000US-0237037P.
				PR	02-OCT-2000;	2000US-0237038P.
				PR	02-OCT-2000;	2000US-0237039P.
				PR	02-OCT-2000;	2000US-0237040P.
				PR	13-OCT-2000;	2000US-0239935P.
				PR	13-OCT-2000;	2000US-0239937P.
				PR	20-OCT-2000;	2000US-0240960P.
				PR	20-OCT-2000;	2000US-0241221P.
				PR	20-OCT-2000;	2000US-0241785P.
				PR	20-OCT-2000;	2000US-0241786P.
				PR	20-OCT-2000;	2000US-0241787P.
				PR	20-OCT-2000;	2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246529P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250319P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465570/50.  
XX Isolated nucleic acid molecule encoding a reproductive system antigen is  
PT used in preventing, treating or ameliorating a medical condition.  
XX  
XX Disclosure; SEQ ID NO 8070; 1297pp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention  
XX  
XX Sequence 6107 BP; 1636 A; 1200 C; 1285 G; 1986 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 1274; DB 4; Length 6107;  
Best Local Similarity 100.0%; Pred. NO. 0;

	Matches 1274;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	TTTCTTTAGTGTGTTATGTCATTTTCTATTTTAGCATTTATTTCTATGTTAGTCTAT	60						
DB	3164	TTTCTTTAGTGTGTTATGTCATTTTCTATTTTAGCATTTATTTCTATGTTAGTCTAT	3223						
QY	61	CCAAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC	120						
DB	3224	CCAAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC	3283						
QY	121	CAGTGTACAGATCCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT	180						
DB	3284	CAGTGTACAGATCCTAATAAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT	3343						
QY	181	TTAATCCTATTTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAACTTTTAAAGGACT	240						
DB	3344	TTAATCCTATTTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAACTTTTAAAGGACT	3403						
QY	241	TATGGCATGTAAACATTTATTTAAGTAAGTCATGTTTATTAATTTTCTCTGCCT	300						
DB	3404	TATGGCATGTAAACATTTATTTAAGTAAGTCATGTTTATTAATTTTCTCTGCCT	3463						
QY	301	CCTTATGTTATTTTTCAGAAATGAGCGTTCAGACTGTACAGCATTTACCTACCGGTACC	360						
DB	3464	CCTTATGTTATTTTTCAGAAATGAGCGTTCAGACTGTACAGCATTTACCTACCGGTACC	3523						
QY	361	TCGAAGTGTCCCAATCCAGAGGGTCTGCTGCTGACTGGCACTGCGCATCCACAAT	420						
DB	3524	TCGAAGTGTCCCAATCCAGAGGGTCTGCTGCTGACTGGCACTGCGCATCCACAAT	3583						
QY	421	GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTTGTAGTGTACGCTGCTTCTT	480						
DB	3584	GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTTGTAGTGTACGCTGCTTCTT	3643						
QY	481	CAATGTACAGTGGCCATCTGTTTGTAGCAACTGTCGCCCAAGCTCACATGTTGCCA	540						
DB	3644	CAATGTACAGTGGCCATCTGTTTGTAGCAACTGTCGCCCAAGCTCACATGTTGCCA	3703						
QY	541	ACTTGGCGGGGCTTTTGGGATCCATTCGCAACTTTGGCTATGGAGAAAGTGGCTAATCA	600						
DB	3704	ACTTGGCGGGGCTTTTGGGATCCATTCGCAACTTTGGCTATGGAGAAAGTGGCTAATCA	3763						
QY	601	GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTAATACTCTGCGCACACAGAA	660						
DB	3764	GTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTAATACTCTGCGCACACAGAA	3823						
QY	661	AAAGCAGACCATGAAGAGCTCTGAGTGTAGGCTTATTCCTGTCGCTGCTGCTGCT	720						
DB	3824	AAAGCAGACCATGAAGAGCTCTGAGTGTAGGCTTATTCCTGTCGCTGCTGCTGCT	3883						
QY	721	TCCTGTAAATGGCAAGGCTCTCTGGATGCTGTATGCCCCCATCTGATGCATCAGCATAAG	780						
DB	3884	TCCTGTAAATGGCAAGGCTCTCTGGATGCTGTATGCCCCCATCTGATGCATCAGCATAAG	3943						
QY	781	TCCATTACAACTCAGGAGAGGATATAGTTTTTCTGCTACAGCATTAATCTTCT	840						
DB	3944	TCCATTACAACTCAGGAGAGGATATAGTTTTTCTGCTACAGCATTAATCTTCT	4003						
QY	841	GGTGTGTTGATGGGTGATGATGAGTCTGCTTTTGGCTTTTCACTTCACTGTTAGTCTTA	900						
DB	4004	GGTGTGTTGATGGGTGATGATGAGTCTGCTTTTGGCTTTTCACTTCACTGTTAGTCTTA	4063						
QY	901	GAGAAACAGGAAATAACGATGTCACAGCAGTGTCTTCGCAATCGTACAGCTGATAGA	960						
DB	4064	GAGAAACAGGAAATAACGATGTCACAGCAGTGTCTTCGCAATCGTACAGCTGATAGA	4123						
QY	961	ACACGCAAGAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGTTCATAGGCGACGA	1020						
DB	4124	ACACGCAAGAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGTTCATAGGCGACGA	4183						
QY	1021	TTGACTTGGGAGCGACTCTCGATCTATTTCATGAGGAAATTCACACAGCCATTAAT	1080						
DB	4184	TTGACTTGGGAGCGACTCTCGATCTATTTCATGAGGAAATTCACACAGCCATTAAT	4243						

QY 1081 AGCGACTGTCAGTCTTTGACACACAGCATTTGACAGCTTTTGCAGAAAATGGCAATTTA 1140  
Db 4244 AGCGACTGTCAGTCTTTGACACACAGCATTTGACAGCTTTTGCAGAAAATGGCAATTTA 4303  
QY 1141 GGCATCAATGTAACATATTTCCATGTTGAAATGGCAATCAACATTTCTGGCCAGTGT 1200  
Db 4304 GGCATCAATGTAACATATTTCCATGTTGAAATGGCAATCAACATTTCTGGCCAGTGT 4363  
QY 1201 TTAAACTTCAGTTTCACAGAAATTAAGGCACCCATCTGTCGCCAAGTAAACTCTTT 1260  
Db 4364 TTAAACTTCAGTTTCACAGAAATTAAGGCACCCATCTGTCGCCAAGTAAACTCTTT 4423  
QY 1261 CGGTAGTGGGAAGC 1274  
Db 4424 CGGTAGTGGGAAGC 4437

RESULT 6

ID ABL98246 standard; DNA; 6107 BP.

XX ABL98246;

21-JUN-2002 (first entry)

DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2898.

Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
reproductive system disorder; urinary system disorder; gene therapy;  
cardiovascular disorder; respiratory disorder; neurological disorder;  
gastrointestinal disease; infection; cytostatic; gene; ds.

OS Homo sapiens.

XX WO200155317-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001329.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189374P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216847P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 06-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239336P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241783P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251130P.  
PR 05-DEC-2000; 2000US-0251198P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483232/52.  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful  
XX for preventing, diagnosing and/or treating testicular cancer.  
XX  
XX Disclosure; SEQ ID NO 2898; 766pp; English.  
XX  
XX The present invention provides the protein and coding sequences of 973  
XX human testicular antigens, and fragments of their genomic sequences. The  
XX sequences can be used in the treatment of cardiovascular, urinary system,  
XX reproductive system, immune, respiratory, neurological and  
XX gastrointestinal disorders, infections, and particularly cancer,  
XX especially testicular cancers. The present sequence is a DNA encoding a  
XX protein fragment of the invention  
XX  
XX SQ Sequence 6107 BP; 1636 A; 1200 C; 1285 G; 1986 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 1274; DB 4; Length 6107;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 TTTCTTTAGTTGTTATGGTCCATTTTCTATTTTATAGCATTTATTTCTATGATGCTAT 60  
XX  
XX 3164 TTTCTTTAGTTGTTATGGTCCATTTTCTATTTTATAGCATTTATTTCTATGATGCTAT 3223  
XX  
XX 61 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAAACATTTTGAAGAGAGAGCTTATC 120  
XX  
XX 3224 CCAAGACGATTAAAGGAGTTCACATGTTTCCGGAAACATTTTGAAGAGAGAGCTTATC 3283  
XX  
XX 121 CAGTGACAGATCTTAATAAAGTGACATTCAGTGAATTTTATTTTATATATCTTTT 180  
XX  
XX 3284 CAGTGACAGATCTTAATAAAGTGACATTCAGTGAATTTTATTTTATATATCTTTT 3343  
XX  
XX 181 TTAATCCTATTTTCTTCTCTTTTCTCAGTAAATTTTGTATGAACCTTAAAGGACT 240  
XX  
XX 3344 TTAATCCTATTTTCTTCTCTTTTCTCAGTAAATTTTGTATGAACCTTAAAGGACT 3403  
XX

QY 241 TATGGCATGTAAACATTTATTTATAAAGTAAGTCATCGTTATTAATTTATTTTCTCTGCCT 300  
DB 3404 TATGGCATGTAAACATTTATTTATAAAGTAAGTCATCGTTATTAATTTATTTTCTCTGCCT 3463  
QY 301 CCTTATGTATTTATTTTACAGAAATGAGCGGTACAGCTGTACAGCATTTACCTACCGGTACC 360  
DB 3464 CCTTATGTATTTATTTTACAGAAATGAGCGGTACAGCTGTACAGCATTTACCTACCGGTACC 3523  
QY 361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGGCCCTGACTGGCACAACACTGCATCCACAAT 420  
DB 3524 TCGAAGTGTCCACCATCCAGAGGGTGCCTGGCCCTGACTGGCACAACACTGCATCCACAAT 3583  
QY 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTGACTATGTGTGTACCGCCATCTTT 480  
DB 3584 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTGACTATGTGTGTACCGCCATCTTT 3643  
QY 481 CAATGTACAGTGGCCATCTTTGTTGTAGCAACTGTGCGCCAAAGCTCACATCTTTGTCCA 540  
DB 3644 CAATGTACAGTGGCCATCTTTGTTGTAGCAACTGTGCGCCAAAGCTCACATCTTTGTCCA 3703  
QY 541 ACTTGGCGGGGCCCTTTGGGATCCATTTCGCAACTTGGCTATGGAGAAAGTGGCTAATTC 600  
DB 3704 ACTTGGCGGGGCCCTTTGGGATCCATTTCGCAACTTGGCTATGGAGAAAGTGGCTAATTC 3763  
QY 601 GTACTTTTCCCTGTAAATATATGCGTCTTCTGATGTGAATAAATCTCTGCCACACAGAA 660  
DB 3764 GTACTTTTCCCTGTAAATATATGCGTCTTCTGATGTGAATAAATCTCTGCCACACAGAA 3823  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTGTAGGCTTATTCCTGTCGCTCGCTGCTGCT 720  
DB 3824 AAAGCAGACCATGAAGAGCTCTGTGAGTGTAGGCTTATTCCTGTCGCTCGCTGCTGCT 3883  
QY 721 TCCTGTAAATGCAAGGCTCTCTGATGTGTAAATGCCCATCTGATGTACATCAGCATAAG 780  
DB 3884 TCCTGTAAATGCAAGGCTCTCTGATGTGTAAATGCCCATCTGATGTACATCAGCATAAG 3943  
QY 781 TCATTAACAACCTCAGGAGAGATATAGTTTTTCTTGTCTACAGACATTAATCTTCTCT 840  
DB 3944 TCATTAACAACCTCAGGAGAGATATAGTTTTTCTTGTCTACAGACATTAATCTTCTCT 4003  
QY 841 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTTGGCTTTTTCACATTTTCTGTTAGTCTTA 900  
DB 4004 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTTGGCTTTTTCACATTTTCTGTTAGTCTTA 4063  
QY 901 GAGAAACAGGAAAAATACGATGTGTCAACGAGCTTCTTCGCAATCGTACAGCTGATAGGA 960  
DB 4064 GAGAAACAGGAAAAATACGATGTGTCAACGAGCTTCTTCGCAATCGTACAGCTGATAGGA 4123  
QY 961 ACAGGACAGCAAGCTGAAAAATTTTGTCTTACCGACTTGAAGTAAATGCTATAGGCGACGA 1020  
DB 4124 ACAGGACAGCAAGCTGAAAAATTTTGTCTTACCGACTTGAAGTAAATGCTATAGGCGACGA 4183  
QY 1021 TTGACTTTGGGAAGGACTCTCGATCTATTTCATGAAGGAATTCACAGCCATTTATGAAT 1080  
DB 4184 TTGACTTTGGGAAGGACTCTCGATCTATTTCATGAAGGAATTCACAGCCATTTATGAAT 4243  
QY 1081 AGCGACTGTCTAGTCTTTTGACACCCAGCATTTGACAGCTTTTTCAGAAAAATGCAATTTA 1140  
DB 4244 AGCGACTGTCTAGTCTTTTGACACCCAGCATTTGACAGCTTTTTCAGAAAAATGCAATTTA 4303  
QY 1141 GGATCAATGTAACTATTTTCCATGTGTGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1200  
DB 4304 GGATCAATGTAACTATTTTCCATGTGTGAATGGAATGGAATGGAATGGAATGGAATGGAAT 4363  
QY 1201 TTAAACTTTCAGTTTTCAGAAAAATGAAGCACTCTCTGCGCACTTAACTCTTTT 1260  
DB 4364 TTAAACTTTCAGTTTTCAGAAAAATGAAGCACTCTCTGCGCACTTAACTCTTTT 4423  
QY 1261 CGGTAGTGGGAAGC 1274  
DB 4424 CGGTAGTGGGAAGC 4437



XX ADS34362;  
XX 02-DEC-2004 (first entry)  
XX POSH protein associated DNA #116.  
XX ds; gene, cytosolic; neurotropic; neuroprotective; antiparkinsonian;  
XX anticonvulsant; antiviral; neuroleptic; central nervous system;  
XX POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI1;  
XX ubiquitin ligase; anti-viral agent; anti-apoptotic agent;  
XX anti-cancer agent; secretory pathway trafficking inhibitor;  
XX neurological disorder progression disorder; Alzheimer's disease;  
XX Parkinson's disease; Huntington's disease; schizophrenia;  
XX Niemann-Pick's disease.  
XX Homo sapiens.  
XX WO20040781130-A2  
XX 16-SEP-2004.  
XX 02-MAR-2004; 2004WO-US006308.  
XX 03-MAR-2003; 2003US-0451437P.  
XX 05-MAR-2003; 2003US-0452284P.  
XX 19-MAR-2003; 2003US-0455760P.  
XX 20-MAR-2003; 2003US-0456640P.  
XX 03-APR-2003; 2003US-0460526P.  
XX 04-APR-2003; 2003US-0460792P.  
XX 21-APR-2003; 2003US-0464285P.  
XX 09-MAY-2003; 2003US-0469462P.  
XX 15-MAY-2003; 2003US-0471378P.  
XX 20-MAY-2003; 2003US-0472327P.  
XX 30-MAY-2003; 2003US-0474706P.  
XX 03-JUN-2003; 2003US-0475925P.  
XX 17-JUN-2003; 2003US-0477317P.  
XX 19-JUN-2003; 2003US-0480915P.  
XX 19-JUN-2003; 2003US-0480376P.  
XX 08-AUG-2003; 2003US-0493850P.  
XX 28-AUG-2003; 2003US-0498639P.  
XX 16-SEP-2003; 2003US-0503931P.  
XX 10-NOV-2003; 2003WO-US035711.  
XX 05-FEB-2004; 2004WO-US003600.  
XX 02-MAR-2004; 2004US-0549896P.  
XX (PROT-) PROTEOLOGICS INC.  
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
XX Greener T;  
XX WPI; 2004-662346/64.  
XX Isolated, purified or recombinant complex, useful for identifying an  
XX antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
XX POSH-associated protein (POSH-AP).  
XX Disclosure; SEQ ID NO 126; 374pp; English.  
XX The invention relates to an isolated, purified or recombinant complex (I)  
XX comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
XX or HERPUDI1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are  
XX useful for identifying an agent that modulates an activity of a POSH  
XX polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
XX apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
XX of a protein through the secretory pathway, an agent that inhibits the  
XX progression of a neurological disorder, an agent that modulates a POSH  
XX function, an agent that modulates a HERPUDI1 function. The methods can be  
XX used for treating a viral infection, for inhibiting an activity of a POSH  
XX AP in a cell, for treating a POSH-associated disease in a subject. The  
XX POSH-associated disease is viral infection, POSH-associated cancer or  
XX POSH-associated neurological disorder. The methods are useful for  
XX treating or preventing POSH-associated neurological disorder in a subject

CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
XX nucleic acid of the invention.  
XX Sequence 2829 BP; 756 A; 566 C; 553 G; 954 T; 0 U; 0 Other;  
Query Match 96.0%; Score 1223; DB 13; Length 2829;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTTCTTAGTGTGTTATGTCCTCAATTTCTATTTTATAGCATTTATTTCTATGAGTCTAT 60  
DB 1134 TTTCTTAGTGTGTTATGTCCTCAATTTCTATTTTATAGCATTTATTTCTATGAGTCTAT 1193  
QY 61 CCAAGACGATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120  
DB 1194 CCAAGACGATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 1253  
QY 121 GAGGTACAGATCCTAATAAGTGCACATTCAGTGAATTTTATTTTATATCTTTT 180  
DB 1254 CAGGTACAGATCCTAATAAGTGCACATTCAGTGAATTTTATTTTATATCTTTT 1313  
QY 181 TTAATCTTATTTTCTCTCTCTTTTCTCAGTAAATTTTGTATGAACCTTTAAAGGACT 240  
DB 1314 TTAATCTTATTTTCTCTCTCTTTTCTCAGTAAATTTTGTATGAACCTTTAAAGGACT 1373  
QY 241 TATGCATGTAACATTTATTTAAGTAAGTCAATGTTTATTAATTTTCTCTGCTCT 300  
DB 1374 TATGCATGTAACATTTATTTAAGTAAGTCAATGTTTATTAATTTTCTCTGCTCT 1433  
QY 301 CCTTATGTTATTTTTCAGAAATGAGCCGTGACATGCTGCTACAGCATTTACTACCGTACC 360  
DB 1434 CCTTATGTTATTTTTCAGAAATGAGCCGTGACATGCTGCTACAGCATTTACTACCGTACC 1493  
QY 361 TCGAAGTGTCCACCATCCAGAGGGTGTCTGCCCTGACTGGCAGCACTGCATCCACAAT 420  
DB 1494 TCGAAGTGTCCACCATCCAGAGGGTGTCTGCCCTGACTGGCAGCACTGCATCCACAAT 1553  
QY 421 GACTTGGCGAGTCTTTTGTGAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCATCTT 480  
DB 1554 GACTTGGCGAGTCTTTTGTGAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCATCTT 1613  
QY 481 CAATGTACAGTGGCCATCTTGTGTTAGCAACCTGTCGCCAAAGTCTCATGTTGTCCA 540  
DB 1614 CAATGTACAGTGGCCATCTTGTGTTAGCAACCTGTCGCCAAAGTCTCATGTTGTCCA 1673  
QY 541 ACTTGGCGGGCGCTTTGGGATCCATTCGCACTTTGGCTATGGAGAACTGCTAATTC 600  
DB 1674 ACTTGGCGGGCGCTTTGGGATCCATTCGCACTTTGGCTATGGAGAACTGCTAATTC 1733  
QY 601 GTACTTTTCCCTGTAAATATATGCTCTTCTGGATGTGAATTAATCTGCGCACACAGAA 660  
DB 1734 GTACTTTTCCCTGTAAATATATGCTCTTCTGGATGTGAATTAATCTGCGCACACAGAA 1793  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTACGCTTATCTCTGCGGCTGCTGCTGCT 720  
DB 1794 AAAGCAGACCATGAAGAGCTCTGTGAGTTTACGCTTATCTCTGCGGCTGCTGCTGCT 1853  
QY 721 TCCTGTAATGCGCAGGCTCTCTGATGCTGTAATGCCCTCATGATGATCAGCATTAAG 780  
DB 1854 TCCTGTAATGCGCAGGCTCTCTGATGCTGTAATGCCCTCATGATGATCAGCATTAAG 1913  
QY 781 TCCTGTAATGCGCAGGCTCTCTGATGCTGTAATGCCCTCATGATGATCAGCATTAAG 840  
DB 1914 TCCTGTAATGCGCAGGCTCTCTGATGCTGTAATGCCCTCATGATGATCAGCATTAAG 1973  
QY 841 GGTGCTGTGAGTGGGTGATGATGAGTCTCTGCTTTTGGCTTTTCACTTCAATGATGCTTA 900  
DB 1974 GGTGCTGTGAGTGGGTGATGATGAGTCTCTGCTTTTGGCTTTTCACTTCAATGATGCTTA 2033  
QY 901 GAGAAACAGGAAATATGATGCTTCCAGGAGTCTTCCGCAATGCTACAGCTGATAGA 960  
DB 2034 GAGAAACAGGAAATATGATGCTTCCAGGAGTCTTCCGCAATGCTACAGCTGATAGA 2093









Db 887 CACAGAAAAGCAGACACCATGAGAGCTCTGTGAGTTTAGGCTTATTCTCTGCGTGCCC 946  
Qy 714 TGGTGCTTCTGTAATAGCAAGGCTCTCTGAGTCTGTAAATGCCCCATCTGATGCATCA 773  
Db 947 TGGTGCTTCTGTAATAGCAAGGCTCTCTGAGTCTGTAAATGCCCCATCTGATGCATCA 1006  
Qy 774 GCATAGTCCATTACACCTACAGGAGAGATAGTTTCTGCTACAGACATTAA 833  
Db 1007 GCATAGTCCATTACACCTACAGGAGAGATAGTTTCTGCTACAGACATTAA 1066  
Qy 834 TCTTCTGCTGCTGTGATGCTGGTGTATGATGAGTCTCTGTTTGGCTTTTCACTTCACTT 893  
Db 1067 TCTTCTGCTGCTGTGATGCTGGTGTATGATGAGTCTCTGTTTGGCTTTTCACTTCACTT 1126  
Qy 894 ACTCTTAGAGAACAGGAAAAATACATGCTCACCAGAGTCTTTCGCAATGTCACACT 953  
Db 1127 ACTCTTAGAGAACAGGAAAAATACATGCTCACCAGAGTCTTTCGCAATGTCACACT 1186  
Qy 954 GATAGAACACGACGAGCTGAAAATTTTCTTACGAGTCTGAGCTAAATGGTCTATAG 1013  
Db 1187 GATAGAACACGACGAGCTGAAAATTTTCTTACGAGTCTGAGCTAAATGGTCTATAG 1246  
Qy 1014 GCGACGATTGACTTGGGAGCGACTCTCTGATCTATTATGAGAAATTCGAACAGCCAT 1073  
Db 1247 GCGACGATTGACTTGGGAGCGACTCTCTGATCTATTATGAGAAATTCGAACAGCCAT 1306  
Qy 1074 TATGAATAGGACTGTCTAGTCTTTGACACGACATTCACAGCTTTTTCGCAAAATGG 1133  
Db 1307 TATGAATAGGACTGTCTAGTCTTTGACACGACATTCACAGCTTTTTCGCAAAATGG 1366  
Qy 1134 CAAATTTAGCATCAATGTAATTTTCCATGTTTGAAATGCAATCAAACTTTTCTGG 1193  
Db 1367 CAAATTTAGCATCAATGTAATTTTCCATGTTTGAAATGCAATCAAACTTTTCTGG 1426  
Qy 1194 CCAGTGTTTAAACTTCAGTTTACAGAAATAAGGACCCATCTGTCTGCGCAACTTAA 1253  
Db 1427 CCAGTGTTTAAACTTCAGTTTACAGAAATAAGGACCCATCTGTCTGCGCAACTTAA 1486  
Qy 1254 ACTCTTTCGTTAGTGGAGC 1274  
Db 1487 ACTCTTTCGTTAGTGGAGC 1507

RESULT 11  
ADS34365  
ID ADS34365 standard; DNA; 1540 BP.  
XX  
AC ADS34365;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE POSH protein associated DNA #119.  
XX  
KW ds; gene; cytosolic; neurotropic; neuroprotective; antiparkinsonian;  
KW anticonvulsant; antiviral; neuroleptic; central nervous system;  
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;  
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
KW anti-cancer agent; secretory pathway trafficking inhibitor;  
KW neurological disorder progression disorder; Alzheimer's disease;  
KW Niemann-Pick's disease; Huntington's disease; schizophrenia;  
XX Homo sapiens.  
XX  
PN W02004078130-A2.  
XX  
PD 16-SEP-2004.  
XX  
PF 02-MAR-2004; 2004WO-US006308.  
XX  
PR 03-MAR-2003; 2003US-0451437P.  
PR 05-MAR-2003; 2003US-0452284P.  
PR 19-MAR-2003; 2003US-0455760P.

PR 20-MAR-2003; 2003US-0456640P.  
PR 03-APR-2003; 2003US-0460526P.  
PR 04-APR-2003; 2003US-0460792P.  
PR 21-APR-2003; 2003US-0464285P.  
PR 09-MAY-2003; 2003US-0469462P.  
PR 15-MAY-2003; 2003US-0471378P.  
PR 20-MAY-2003; 2003US-0472327P.  
PR 30-MAY-2003; 2003US-0474706P.  
PR 03-JUN-2003; 2003US-0475825P.  
PR 17-JUN-2003; 2003US-0479317P.  
PR 19-JUN-2003; 2003US-0480215P.  
PR 19-JUN-2003; 2003US-0480376P.  
PR 08-AUG-2003; 2003US-0493860P.  
PR 28-AUG-2003; 2003US-0498634P.  
PR 16-SEP-2003; 2003US-0503931P.  
PR 10-NOV-2003; 2003WO-US035712.  
PR 05-FEB-2004; 2004WO-US003600.  
PR 02-MAR-2004; 2004US-0549896P.  
XX  
XX (PROT-) PROTEOLOGICS INC.  
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
XX Greener T;  
XX WPI; 2004-662346/64  
XX Isolated, purified or recombinant complex, useful for identifying an  
XX antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
XX POSH-associated protein (POSH-AP).  
XX Disclosure; SEQ ID NO 129; 374pp; English.  
XX The invention relates to an isolated, purified or recombinant complex (I)  
XX comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
XX or HERPUD1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are  
XX useful for identifying an agent that modulates an activity of a POSH  
XX polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
XX apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
XX of a protein through the secretory pathway, an agent that modulates a POSH  
XX progression of a neurological disorder, an agent that modulates a POSH  
XX function, an agent that modulates a Herpud1 function. The methods can be  
XX used for treating a viral infection, for inhibiting an activity of a POSH  
XX -AP in a cell, for treating a POSH-associated disease in a subject. The  
XX POSH-associated disease is viral infection, POSH-associated cancer or  
XX POSH-associated neurological disorder. The methods are useful for  
XX treating or preventing POSH-associated neurological disorder in a subject  
XX e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
XX schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
XX nucleic acid of the invention.  
XX  
XX SQ Sequence 1540 BP; 428 A; 313 C; 352 G; 447 T; 0 U; 0 Other;

Query Match 75.2%; Score 958; DB 13; Length 1540;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 317 CAGAAATGAGCGCTCAGACTGTACAGATTCACCTACCGTACCTCGAGTCTCCACCAT 376  
Db 113 CAGAAATGAGCGCTCAGACTGTACAGATTCACCTACCGTACCTCGAGTCTCCACCAT 172  
Qy 377 CCAGAGGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436  
Db 173 CCAGAGGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 232  
Qy 437 TTGAGTGTCCAGTCTGCTTGTGCTATGTTTACCGCCCATTTCTCAATGTGAGTGGCC 496  
Db 233 TTGAGTGTCCAGTCTGCTTGTGCTATGTTTACCGCCCATTTCTCAATGTGAGTGGCC 292  
Qy 497 ATCTTGTGTTAGCAACTGTGCGCCAAAGCTCAGATGTTGTCCAACTTCGCGGCGCTT 556  
Db 293 ATCTTGTGTTAGCAACTGTGCGCCAAAGCTCAGATGTTGTCCAACTTCGCGGCGCTT 352  
Qy 557 TGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTGA 616











Db 3191 GACTGGGAGTCTTTTGGAGTGCAGTCTGTTGACTATGTGTACGCCCATCTTT 3132  
Qy 481 CAATGTGAGAGTGGCCATCTTGTGTAGCAACTGTGCGCCAAAGTCTCATGTCTTCCA 540  
Db 3131 CAATGTGAGAGTGGCCATCTTGTGTAGCAACTGTGCGCCAAAGTCTCATGTCTTCCA 3072  
Qy 541 ACTTGGCGGGGCGCTTGGATCCATTCGCAACTGTGCGCTATGAGAAAGTGGCTAATCCA 600  
Db 3071 ACTTGGCGGGGCGCTTGGATCCATTCGCAACTGTGCGCTATGAGAAAGTGGCTAATCCA 3012  
Qy 601 GTACTTTTCCCTGTAATATGCTCTCTGGATGTGAATACTCTGCCACACAGAA 660  
Db 3011 GTACTTTTCCCTGTAATATGCTCTCTGGATGTGAATACTCTGCCACACAGAA 2952  
Qy 661 AAA-GCAGACCAATGAAGA-GCTCTGTGAGTTTAGGCCCTTAATCTCTGCGTGCCTTGGTG 718  
Db 2951 AAANGCAGACCAATGAANGCTCTGTGAGTTTAGGCCCTTAATCTCTGCGTGCCTTGGTG 2892  
Qy 719 CTTCTGTAAATGGCAAGGCTCTCTGGATGTCTGTAATGCCCATCTGATGCATCAGCATA 778  
Db 2891 CTTCTGTAAATGGCAAGGCTCTCTGGATGTCTGTAATGCCCATCTGATGCATCAGCATA 2832  
Qy 779 AGTCCATTACAACTTACAGGAGAGATATAGTTTCTTGTCTACAGACATTAATCTTC 838  
Db 2831 AGTCCATTACAACTTACAGGAGAGATATAGTTTCTTGTCTACAGACATTAATCTTC 2772  
Qy 839 CTGGTCTGTGACTGGGTGATGATCAGTCTCTGTTTGTGCTTCTTCACTTGTAGTCT 898  
Db 2771 CTGGTCTGTGACTGGGTGATGATCAGTCTCTGTTTGTGCTTCTTCACTTGTAGTCT 2712  
Qy 899 TAGAGAAAACAGAAATACGATGTTTACAGCAGTCTTTCGCAATCGTACAGCTGATAG 958  
Db 2711 TAGAGAAAACAGAAATACGATGTTTACAGCAGTCTTTCGCAATCGTACAGCTGATAG 2652  
Qy 959 GAACGCGCAAGCAAGCTGAAATTTTCTTACCGACTTCAGCTTAATGTCTATAGCGGAC 1018  
Db 2651 GAACGCGCAAGCAAGCTGAAATTTTCTTACCGACTTCAGCTTAATGTCTATAGCGGAC 2592  
Qy 1019 GATTGACTTTGGGAAGCGACTCTCTCGATCTATTATGAAAGAAATGCAACAGCCATTATGA 1078  
Db 2591 GATTGACTTTGGGAAGCGACTCTCTCGATCTATTATGAAAGAAATGCAACAGCCATTATGA 2532  
Qy 1079 ATAGCGACTGTCTAGTCTTTACACAGCAGATTCACAGCTTTTTCAGAAATGCGAATT 1138  
Db 2531 ATAGCGACTGTCTAGTCTTTACACAGCAGATTCACAGCTTTTTCAGAAATGCGAATT 2472  
Qy 1139 TAGGCATCAATGTAATCTTCCATGTGTGAAATGCAATCAAACTTTCTGGCCAGT 1198  
Db 2471 TAGGCATCAATGTAATCTTCCATGTGTGAAATGCAATCAAACTTTCTGGCCAGT 2412  
Qy 1199 GTTTAAACTTCAGTTTACAGAAATAAGGCAACCCATCTGTGCAACCTTAAACTCT 1258  
Db 2411 GTTTAAACTTCAGTTTACAGAAATAAGGCAACCCATCTGTGCAACCTTAAACTCT 2352  
Qy 1259 TTTCGTPAGTGAAGC 1274  
Db 2351 TTTCGTPAGTGAAGC 2336

## RESULT 16

ADS34364

ID ADS34364 standard; DNA; 2034 BP.

AC ADS34364;

XX ADS34364;

DT 02-DEC-2004 (first entry)

XX POSH protein associated DNA #118.

DE ds; gene; cytostatic; neurotropic; neuroprotective; antiparkinsonian;

KW anticonvulsant; antiviral; neuroleptic; central nervous system;

KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;

XX

KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
KW anti-cancer agent; secretory pathway trafficking inhibitor;  
KW neurological disorder progression disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; schizophrenia;  
KW Niemann-Pick's disease.  
XX Homo sapiens.  
OS  
XX WO2004078130-A2.  
XX 16-SEP-2004.  
XX 02-MAR-2004; 2004WO-US006308.  
XX 03-MAR-2003; 2003US-0451437P.  
XX 05-MAR-2003; 2003US-0452284P.  
XX 19-MAR-2003; 2003US-0455760P.  
XX 20-MAR-2003; 2003US-0456640P.  
XX 03-APR-2003; 2003US-0460526P.  
XX 04-APR-2003; 2003US-0460792P.  
XX 21-APR-2003; 2003US-0464285P.  
XX 09-MAY-2003; 2003US-0469462P.  
XX 15-MAY-2003; 2003US-0471378P.  
XX 20-MAY-2003; 2003US-0472327P.  
XX 30-MAY-2003; 2003US-0474706P.  
XX 03-JUN-2003; 2003US-0475825P.  
XX 17-JUN-2003; 2003US-0479317P.  
XX 19-JUN-2003; 2003US-0480215P.  
XX 19-JUN-2003; 2003US-0480376P.  
XX 08-AUG-2003; 2003US-0493860P.  
XX 28-AUG-2003; 2003US-0498634P.  
XX 16-SEP-2003; 2003US-0503931P.  
XX 10-NOV-2003; 2003WO-US035712.  
XX 05-FEB-2004; 2004WO-US003600.  
XX 02-MAR-2004; 2004US-0549896P.  
XX  
XX (PROT-) PROTEOLOGICS INC.  
XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
XX Greener T;  
XX WPI; 2004-662346/64.

Isolated, purified or recombinant complex, useful for identifying an antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and POSH-associated protein (POSH-AP).

Disclosure; SEQ ID NO 128A 374pp; English.

The invention relates to an isolated, purified or recombinant complex (I) comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a) or HERPUD1 and a ubiquitin ligase (b). Methods using (I), (a) or (b) are useful for identifying an agent that modulates an activity of a POSH polypeptide or POSH-AP, for identifying an antiviral agent, an anti-apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking of a protein through the secretory pathway, an agent that inhibits the progression of a neurological disorder, an agent that modulates a POSH function, an agent that modulates a HERPUD1 function. The methods can be used for treating a viral infection, for inhibiting an activity of a POSH-AP in a cell, for treating a POSH-associated disease in a subject. The POSH-associated disease is viral infection, POSH-associated cancer or POSH-associated neurological disorder. The methods are useful for treating or preventing POSH-associated neurological disorder in a subject e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease, schizophrenia, Niemann-Pick's disease. This sequence corresponds to a nucleic acid of the invention.

XX SQ Sequence 2034 BP; 580 A; 378 C; 424 G; 652 T; 0 U; 0 Other;

Query Match

Best Local Similarity 65.8%; Score 838; DB 13; Length 2034;

Matches 958; Conservative 0; Mismatches 0; Indels 1; Gaps 1;



317	QY	CAGAAATGAGCGT	CAGACTGCTACAGCATTA	CTTACCGGTACCTCGAAGTGTCCACCAAT	376
100	Db	CAGAAATGAGCGGT	CAGACTGCTTACAGCATTA	CTTACCGGTACCTCGAAGTGTCCACCAAT	159
377	QY	CCGAGGGTGGCTG	CCCTGCACTGGGCACAACT	GCATCCAACTGACTGGGGGAGTCTTTT	436
160	Db	CCGAGGGTGGCTG	CCCTGCACTGGGCACAACT	GCATCCAACTGACTGGGGGAGTCTTTT	219
437	QY	TTGAGTGTCCAGT	CTGCTTTGACTATGTGTAC	CCGCCCAATCTTCAATGTCAAGTGGCC	496
220	Db	TTGAGTGTCCAGT	CTGCTTTGACTATGTGTAC	CCGCCCAATCTTCAATGTCAAGTGGCC	279
497	QY	ATCTGTTTGTAG	CAACTGTGCGCCAAAGCT	CACTGTTGTCCAACTGCGGGGGCCCTT	556
280	Db	ATCTGTTTGTAG	CAACTGTGCGCCAAAGCT	CACTGTTGTCCAACTGCGGGGGCCCTT	339
557	QY	TGGGATCCATT	CGCAACTTTGGCTATGG	AAAGTGGCTAATTCAGTACTTTTCCCTGTGA	616
340	Db	TGGGATCCATT	CGCAACTTTGGCTATGG	AAAGTGGCTAATTCAGTACTTTTCCCTGTGA	399
617	QY	AATATCGGTCTT	CTGGATGTGAAATAACT	CTGCCACACAGAAAAAGCAGACCATAAG	676
400	Db	AATATCGGTCTT	CTGGATGTGAAATAACT	CTGCCACACAGAAAAAGCAGACCATAAG	459
677	QY	AGCTCTGTGAGT	TTAGGCCCTTATTCCTG	TCGTCGCTCGTCTCTGTAAATGCAAG	736
460	Db	AGCTCTGTGAGT	TTAGGCCCTTATTCCTG	TCGTCGCTCGTCTCTGTAAATGCAAG	519
737	QY	GCTCTCTGATG	CTGTAAATGCCCCCAT	CTGATGCATCAGCATTAAGTCCATTACAACTTAC	796
520	Db	GCTCTCTGATG	CTGTAAATGCCCCCAT	CTGATGCATCAGCATTAAGTCCATTACAACTTAC	579
797	QY	AGGGAGGAGAT	TAGTTTTTCTTGCTAC	GACAACTTAATCTTCTGCTGCTGTTGACTGG	856
580	Db	AGGGAGGAGAT	TAGTTTTTCTTGCTAC	GACAACTTAATCTTCTGCTGCTGTTGACTGG	639
857	QY	TGATGATG	CAGTCCCTGTTTGCTT	TCATTCATGTTAGTCTTTAGAGAAACAGGAAAAAT	916
640	Db	TGATGATG	CAGTCCCTGTTTGCTT	TCATTCATGTTAGTCTTTAGAGAAACAGGAAAAAT	699
917	QY	ACGATGGT	CACGACGAGTTCTT	CGCAATCGTACAGCTGATGCAACCCGACCAAGCTG	976
700	Db	ACGATGGT	CACGACGAGTTCTT	CGCAATCGTACAGCTGATGCAACCCGACCAAGCTG	759
977	QY	AAAAATTTGCTT	ACCGACTTGAGCTAAAT	GGTCAATAGGCGCACGATTGACTTGGGAAGCGA	1036
760	Db	AAAAATTTGCTT	ACCGACTTGAGCTAAAT	GGTCAATAGGCGCACGATTGACTTGGGAAGCGA	819
1037	QY	CTCCTCGATCT	ATTCAATGAGGAATT	GCACACGCCAATTATGAATAGCGACTGTCTAGTCT	1096
820	Db	CTCCTCGATCT	ATTCAATGAGGAATT	GCACACGCCAATTATGAATAGCGACTGTCTAGTCT	879
1097	QY	TTGACAC	CAGCATTCGACAGC	-TTTTTGCAGAAATGGCAATTTTAGGCATCAATGTA	1155
880	Db	TTGACAC	CAGCATTCGACAGC	-TTTTTGCAGAAATGGCAATTTTAGGCATCAATGTA	939
1156	QY	ATTTTCCATG	TGTTGAAATGGCAAT	CAAACTTTTCTGGCCAGTGTGTTTAAAACTTCAGTTT	1215
940	Db	ATTTTCCATG	TGTTGAAATGGCAAT	CAAACTTTTCTGGCCAGTGTGTTTAAAACTTCAGTTT	999
1216	QY	CACAGAAAA	TAAGGCCACCCCAT	CTGTGCGCAACTAAAACTTTTTCGGTAGTGGGAAGC	1274
1000	Db	CACAGAAAA	TAAGGCCACCCCAT	CTGTGCGCAACTAAAACTTTTTCGGTAGTGGGAAGC	1058

RESULT 17  
AAT64820

AAT64820  
ID AAT64820 standard; cDNA: 1884 BP.

XX  
AC AAT64820;XX  
DT 25-MAR-1998 (first entry)

XXIX

Tumour suppressor activated pathway gene TSAP3.

Tumour suppressor activated pathway; tumour suppressor inhibited pathway; TSAP; TSP; HUM5IAH; Human Homologue of the *Drosophila* seven in absentia; apoptosis; p3; cancer; Alzheimer's disease; probe; amplification; PCR; primer; hybridisation; diagnosis; ss.

**Homo sapiens.**

Key	Location/Qualifiers
1	1.1
2	2.1
3	3.1
4	4.1
5	5.1
6	6.1
7	7.1
8	8.1
9	9.1
10	10.1
11	11.1
12	12.1
13	13.1
14	14.1
15	15.1
16	16.1
17	17.1
18	18.1
19	19.1
20	20.1
21	21.1
22	22.1
23	23.1
24	24.1
25	25.1
26	26.1
27	27.1
28	28.1
29	29.1
30	30.1
31	31.1
32	32.1
33	33.1
34	34.1
35	35.1
36	36.1
37	37.1
38	38.1
39	39.1
40	40.1
41	41.1
42	42.1
43	43.1
44	44.1
45	45.1
46	46.1
47	47.1
48	48.1
49	49.1
50	50.1
51	51.1
52	52.1
53	53.1
54	54.1
55	55.1
56	56.1
57	57.1
58	58.1
59	59.1
60	60.1
61	61.1
62	62.1
63	63.1
64	64.1
65	65.1
66	66.1
67	67.1
68	68.1
69	69.1
70	70.1
71	71.1
72	72.1
73	73.1
74	74.1
75	75.1
76	76.1
77	77.1
78	78.1
79	79.1
80	80.1
81	81.1
82	82.1
83	83.1
84	84.1
85	85.1
86	86.1
87	87.1
88	88.1
89	89.1
90	90.1
91	91.1
92	92.1
93	93.1
94	94.1
95	95.1
96	96.1
97	97.1
98	98.1
99	99.1
100	100.1

key	1. .807
CDS	

```

/*tag= a
/product= "human TSAP3 "

```

WO9722695-A2.

26-JUN-1997.

20-DEC-1996: 96WO-FR002061.

20-DEC-1995: 95ER-00015146.

20-DEC-1995; 95FR-00013148:  
18-APR-1996: 96FR-00004853:

(DAUS-) FOND DAUSSET-CEPH JEAN.

Tajerman A. Amson R. Cohen D:

WPT: 1997-341686/31.

WPI; 1997-341686/  
P-PSDB: AAW18520;

New genes activated or inhibited during apoptosis - useful for treatment of... and for assessing risk of developing, cancer and Alzheimer's disease.

Claim 1: page 29-30: 51pp: French.

This is the nucleotide sequence of the human tumour suppressor activated pathway gene 3 (TSAP3). The sequences AAT64813-T64823 represent genes of the tumour suppressor activated pathway (TSAP 1-8, also human TSAP 3, also designated HUM5IAH, i.e. Human Homologue of the *Drosophila* seven in absentia gene) or of the tumour suppressor inhibited pathway (TSIP 1 and 2). Expression of TSAP genes is induced during apoptosis while that of TSIP is induced by tumour suppressors and inhibited during apoptosis, especially where apoptosis is induced by p53. The sequences, vectors, containing them or compounds that induce their expression are useful for treatment of cancer and Alzheimer's disease. Fragments of the sequences are useful as probes and amplification primers to determine predisposition to these diseases by detecting abnormalities in the genes.

Sequence 1884 BP: 524 A: 370 C: 372 G: 618 T: 0 U: 0 Other:

61.2%: Score 780: DB 2: Length 1884:

Very Match	81.2%;	Score 780;	Pred. No. 0;
Not Local Similarity.	100.0%;		

Local Similarity: 100.0%, Rec. No.: 0;  
 Mismatches 780; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

322 ATGAGCCGTCAGACTGCTACAGCATTACCTACCGGTACCTCGAAGTGTCACCATCCCAG 381

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

[illegible]

382 AGGGTGCCTGCCCIGACCGCAGCACATCGCATCCATCAATGACCTGGCCAGATCTTTTCATG

61 AGGGTGGCTGGCCTGAC TGGCACAAC TGCATCCATCACA TATGACTTGGCCAGATCTTTTGGAS 120

442 TGTCCAGTCTGCTTTGACTATGTGTACCGCCATTCTCAAGTCAGAGTGGCCATCTT 501

121 TGTCCAGTCTGTCTTGGACTATGTGTACCGCCCAATCTCTCAATGTCAGAGTGGCCATCTT 180

502 GTTGTAGCAACTGTGCGCCCAAGCTCACATGTTGTCCAACTTGCCGGGGCCCTTTGGGA 561

181 GTTGTAGCAACTGTGCGCCCAAAGCTCACATGTTGTCCAACTTGCCGGGGCCCTTTGGGA 240

562 TCCATTTCGCAACTTTGGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCCCTGTAATAT 621



Db 2651 GCTCTCTGGATGCTGTAATGCCCCCATCTGATGCATCAGCATTAAGTCCATTACAAACCCCTAC 2592  
QY 797 AGGAGAGAGATAGATTTTCTTCTGTACAGACATTAATCTTCTGGTGTGTTGACTGGG 856  
Db 2591 AGGAGAGAGATAGATTTTCTTCTGTACAGACATTAATCTTCTGGTGTGTTGACTGGG 2532  
QY 857 TGATGATGAGTCCCTGTTTGGCTTTCACATTCATGTTAGTC 897  
Db 2531 TGATGATGAGTCCCTGTTTGGCTTTCACATTCATGTTAGTC 2491

## RESULT 19

AAC98856  
ID AAC98856 standard; cDNA; 1535 BP.  
XX  
AC AAC98856;  
XX  
DT 09-MAR-2001 (first entry)  
XX  
DE Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:84.  
XX  
KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;  
KW diagnosis; identification; cytostatic; neuroprotective; nootropic;  
KW immunomodulatory; relaxant; contraceptive; gynaecological;  
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;  
KW linkage analysis; tissue identification; tissue typing; forensic; neural;  
KW immune system; muscular; reproductive; gastrointestinal; pulmonary;  
KW cardiovascular; renal; proliferative; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200055320-A1.  
XX  
PD 21-SEP-2000.  
XX  
PF 08-MAR-2000; 2000WO-US005989.  
XX  
PX 12-MAR-1999; 99US-0124270P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM;  
XX  
DR WPI: 2000-579444/54.  
DR P-PSDB; AAB54091.  
XX

New nucleic acid that is a pancreatic cancer antigen for preventing, treating, or ameliorating a medical condition, particular pancreatic cancer, or for use in assays for diagnosing a pathological condition.

Claim 1; Page 554-555; 1379pp; English

AAC98873 to AAC99231 encode the human pancreatic cancer associated proteins, called pancreatic cancer antigens, given in AAB54008 to AAB54466. The human pancreatic cancer antigens have cytostatic, neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive, gynaecological, cardiant and antiinflammatory activities, and can be used in gene therapy. The polynucleotide and proteins can be used for preventing, treating, or ameliorating a medical condition or in assays for diagnosing a pathological condition or a susceptibility to one in a subject. Binding partners to the proteins and the activity of the proteins can be identified. The pancreatic cancer antigens can be used to detect, treat or prevent pancreatic disorders, especially cancer. Agonists and antagonists to the antigens can be screened for. The pancreatic cancer antigen polynucleotides can be used to design nucleic acid hybridisation probes that can be used in chromosome mapping, linkage analysis, tissue identification and/or typing and a variety of forensic and diagnostic methods. The proteins can be used to generate antibodies which are used to purify, detect and target the polypeptides, including both in vivo and in vitro diagnostic and therapeutic methods. The proteins can be used to treat or prevent neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal or

CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent sequences used in the exemplification of the present invention  
XX  
SQ Sequence 1535 BP; 459 A; 243 C; 283 G; 544 T; 0 U; 6 Other;

Query Match 27.9%; Score 356; DB 3; Length 1535;  
Best Local Similarity 100.0%; Pred. No. 5e-166;  
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 GATGTCACACGAGCTTCTTCGCAATCGTACAGCTGATAGGAAACACGCAAGCAAGCTGAA 978  
Db 36 GATGTCACACGAGCTTCTTCGCAATCGTACAGCTGATAGGAAACACGCAAGCAAGCTGAA 95  
QY 979 AATTTTGTACCGACTTTCGCAATCGTACAGCTGATAGGAAACACGCAAGCTGAA 1038  
Db 96 AATTTTGTACCGACTTTCGCAATCGTACAGCTGATAGGAAACACGCAAGCTGAA 155  
QY 1039 CCTCATCTATTTCATGAAGGAATTCGCAAGCAAGCTGATAGGAAACACGCAAGCTGAA 1098  
Db 156 CCTCATCTATTTCATGAAGGAATTCGCAAGCAAGCTGATAGGAAACACGCAAGCTGAA 215  
QY 1099 GACACGAGCTTTCGCAAGCAAGCTGATAGGAAACACGCAAGCTGAA 1158  
Db 216 GACACGAGCTTTCGCAAGCAAGCTGATAGGAAACACGCAAGCTGAA 275  
QY 1159 TCCATGCTTGAATGCGCAATCAACATTTCTGCGCCAGTGTAAAACTTCAGTTTCAC 1218  
Db 276 TCCATGCTTGAATGCGCAATCAACATTTCTGCGCCAGTGTAAAACTTCAGTTTCAC 335  
QY 1219 AGAAAAAAGGACCCCATCTGTCGCAACCTAAACCTTTTCGGTAGTGAAGC 1274  
Db 336 AGAAAAAAGGACCCCATCTGTCGCAACCTAAACCTTTTCGGTAGTGAAGC 391

## RESULT 20

AAB7326/c  
ID AAB7326 standard; cDNA; 466 BP.  
XX  
AC AAB7326;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 7386.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US004927.  
XX  
PR 28-FEB-2000; 2000US-00515126.  
PR 18-MAY-2000; 2000US-00577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI: 2001-514838/56.  
DR P-PSDB; AAO07395.  
XX

Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukemia, inflammation and immune disorders.

Claim 1; SEQ ID NO 7386; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AAB79941-AAB93841) and





CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 473 BP; 122 A; 104 C; 124 G; 123 T; 0 U; 0 Other;

Query Match 10.0%; Score 128; DB 5; Length 473;  
Best Local Similarity 100.0%; Pred. No. 9,2e-53;  
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 969 GCAAGCTGAAATTTTGGCTTACCGACTTGAAGCTTGAATGTCATAGCGGAGTGTG 1028  
DB 473 GCAAGCTGAAATTTTGGCTTACCGACTTGAAGCTTGAATGTCATAGCGGAGTGTG 414  
QY 1029 GGAAGCGACTCTCGATCTATTATCATCAAGCAATTGCAACAGCCATTATGATAGCGACTG 1088  
DB 413 GGAAGCGACTCTCGATCTATTATCATCAAGCAATTGCAACAGCCATTATGATAGCGACTG 354  
QY 1089 TCTAGTCT 1096  
DB 353 TCTAGTCT 346

RESULT 25  
ABV13828/c  
ID ABV13828 standard; cDNA; 411 BP.  
XX  
AC ABV13828;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 13819.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JB;  
XX  
DR WPI; 2001-662795/76.  
XX

Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 2301; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 411 BP; 153 A; 70 C; 74 G; 114 T; 0 U; 0 Other;

Query Match 9.3%; Score 118; DB 5; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8.5e-48;  
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 TTTCATGTGTTGAATGCAATCAACATTTTCTGGCCAGTGTTTAAACTTCAGTTTC 1216  
DB 364 TTTCATGTGTTGAATGCAATCAACATTTTCTGGCCAGTGTTTAAACTTCAGTTTC 305  
QY 1217 ACAGAAATAAGCGCCCATCTGTCTGCCAACCTAAACTCTTCGTTAGTGGAGC 1274  
DB 304 ACAGAAATAAGCGCCCATCTGTCTGCCAACCTAAACTCTTCGTTAGTGGAGC 247

RESULT 26  
ABV34940/c  
ID ABV34940 standard; cDNA; 441 BP.  
XX  
AC ABV34940;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 34931.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JB;  
XX  
DR WPI; 2001-662795/76.  
XX

Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 7303; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 441 BP; 157 A; 80 C; 86 G; 118 T; 0 U; 0 Other;  
 Query Match 9.3%; Score 118; DB 5; Length 441;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-48;  
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1157 TTTCATGTTGTAATGCAATCAACATTTCTGCGCAGTGTTTAAACATTCAGTTTC 1216  
 Db 399 TTTCATGTTGTAATGCAATCAACATTTCTGCGCAGTGTTTAAACATTCAGTTTC 340  
 QY 1217 ACAGAAATAAGGACCCATCTGTCGCAACTAAACTCTTTTCGGTAGTGGAAGC 1274  
 Db 339 ACAGAAATAAGGACCCATCTGTCGCAACTAAACTCTTTTCGGTAGTGGAAGC 282  
 RESULT 27  
 ABV04659/c  
 ID ABV04659 standard; cDNA; 438 BP.  
 AC ABV04659;  
 XX  
 DT 13-SEP-2002 (first entry)  
 DE Human prostate expression marker cDNA 4650.  
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 KW Homo sapiens.  
 OS  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US005171.  
 XX  
 PR 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC  
 XX  
 PI Schlegel R, Endege WO, Monahan JB;  
 XX  
 WPI; 2001-662795/76.  
 XX  
 DR Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1; Page 805; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (i) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (i) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 438 BP; 149 A; 79 C; 85 G; 122 T; 0 U; 3 Other;  
 Query Match 6.7%; Score 85; DB 5; Length 438;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-31;  
 Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1190 CTGGCAGTGTTTAAACTTCAGTTTCACAGAAATAAGGCCACCATCTGTCTGCCAAC 1249  
 Db 366 CTGGCAGTGTTTAAACTTCAGTTTCACAGAAATAAGGCCACCATCTGTCTGCCAAC 307  
 QY 1250 TAAACTCTTTTCGGTAGTGGAAGC 1274  
 Db 306 TAAACTCTTTTCGGTAGTGGAAGC 282  
 RESULT 28  
 ABI99429  
 ID ABI99429 standard; cDNA; 1968 BP.  
 AC ABI99429;  
 XX  
 DT 07-MAR-2002 (first entry)  
 DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:354.  
 DE Mouse; ischaemic; compressive ischaemia; occlusive ischaemia;  
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.  
 XX  
 OS Mus musculus.  
 XX  
 PN WO200188188-A2.  
 XX  
 PD 22-NOV-2001.  
 XX  
 PF 18-MAY-2001; 2001WO-JP004192.  
 XX  
 PR 18-MAY-2000; 2000JP-00145977.  
 XX  
 PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.  
 XX  
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;  
 XX  
 WPI; 2002-034733/04.  
 DR P-PSDB; ABB57146.  
 XX  
 PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
 PT expression levels of particular genes defined in the specification or by  
 PT determining the expression profile of a gene group comprising these  
 PT genes.  
 XX  
 PS Claim 2; Page 956-958; 2690pp; English.  
 XX  
 CC The present invention describes a method for examining ischaemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC genes (i) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (i). The method  
 CC is useful for examining the ischaemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
 CC expression levels of particular genes (AB199202 to AB199912, encoding the  
 CC protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The expression  
 CC levels or expression profiles produced by these genes are used as an  
 CC indicator when screening for ischaemic condition-improving drugs or  
 CC therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR  
 CC primers for a mouse ischaemic condition related sequence, which are used  
 CC in the exemplification of the present invention  
 XX  
 SQ Sequence 1968 BP; 500 A; 397 C; 447 G; 624 T; 0 U; 0 Other;  
 Query Match 6.0%; Score 77; DB 6; Length 1968;  
 Best Local Similarity 100.0%; Pred. No. 2e-27;





Mon Apr 25 11:06:36 2005

us-10-679-246-1.oli15.rng

```
SQ Sequence 426 BP; 105 A; 84 C; 87 G; 109 T; 0 U; 41 Other;
Query Match 4.1%; Score 52; DB 5; Length 426;
Best Local Similarity 100.0%; Pred. No. 5.2e-15;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 965 GCAAGCAGCTGAAATTTGTTACCGACTTGACCTAAATGGTCATAGGCG 1016
Db 426 GCAAGCAGCTGAAATTTGTTACCGACTTGACCTAAATGGTCATAGGCG 375

RESULT 31
ID AAL01548 standard; cDNA; 222 BP.
AC AAL01548;
XX
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen cDNA SEQ ID NO: 1549.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ss.
XX
XX Homo sapiens;
XX WO200155320-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001339.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225366P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225368P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226686P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 08-SEP-2000; 2000US-0231968P.
XX 12-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 25-SEP-2000; 2000US-0234998P.
XX 26-SEP-2000; 2000US-0235484P.
XX 27-SEP-2000; 2000US-0235834P.
XX 27-SEP-2000; 2000US-0235836P.
XX 29-SEP-2000; 2000US-0236327P.
XX 29-SEP-2000; 2000US-0236367P.
XX 29-SEP-2000; 2000US-0236368P.
XX 29-SEP-2000; 2000US-0236369P.
XX 29-SEP-2000; 2000US-0236370P.
XX 02-OCT-2000; 2000US-0236802P.
XX 02-OCT-2000; 2000US-0237037P.
XX 02-OCT-2000; 2000US-0237038P.
XX 02-OCT-2000; 2000US-0237039P.
XX 02-OCT-2000; 2000US-0237040P.
XX 13-OCT-2000; 2000US-0239935P.
XX 13-OCT-2000; 2000US-0239937P.
XX 20-OCT-2000; 2000US-0240960P.
XX 20-OCT-2000; 2000US-0241221P.
XX 20-OCT-2000; 2000US-0241785P.
XX 20-OCT-2000; 2000US-0241786P.
XX 20-OCT-2000; 2000US-0241787P.
XX 20-OCT-2000; 2000US-0241808P.
XX 20-OCT-2000; 2000US-0241809P.
XX 20-OCT-2000; 2000US-0241826P.
XX 01-NOV-2000; 2000US-0244617P.
XX 08-NOV-2000; 2000US-0246474P.
XX 08-NOV-2000; 2000US-0246475P.
XX 08-NOV-2000; 2000US-0246476P.
XX 08-NOV-2000; 2000US-0246477P.
XX 08-NOV-2000; 2000US-0246478P.
XX 08-NOV-2000; 2000US-0246523P.
XX 08-NOV-2000; 2000US-0246524P.
XX 08-NOV-2000; 2000US-0246525P.
XX 08-NOV-2000; 2000US-0246526P.
XX 08-NOV-2000; 2000US-0246527P.
XX 08-NOV-2000; 2000US-0246528P.
XX 08-NOV-2000; 2000US-0246532P.
XX 08-NOV-2000; 2000US-0246609P.
XX 08-NOV-2000; 2000US-0246610P.
XX 08-NOV-2000; 2000US-0246611P.
XX 08-NOV-2000; 2000US-0246613P.
XX 17-NOV-2000; 2000US-0249207P.
XX 17-NOV-2000; 2000US-0249208P.
XX 17-NOV-2000; 2000US-0249209P.
XX 17-NOV-2000; 2000US-0249210P.
XX 17-NOV-2000; 2000US-0249211P.
XX 17-NOV-2000; 2000US-0249212P.
XX 17-NOV-2000; 2000US-0249213P.
XX 17-NOV-2000; 2000US-0249214P.
XX 17-NOV-2000; 2000US-0249215P.
XX 17-NOV-2000; 2000US-0249216P.
```

PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-465570/50.  
DR P-PSDB; AAM95578.  
XX  
PT Isolated nucleic acid molecule encoding a reproductive system antigen is  
PT used in preventing, treating or ameliorating a medical condition.  
XX  
XX Claim 1; SEQ ID NO 1549; 1297pp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a coding sequence of the  
XX invention  
XX  
SQ Sequence 222 BP; 53 A; 52 C; 64 G; 50 T; 0 U; 3 Other;  
Query Match 2.8%; Score 33; DB 4; Length 222;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 318 AGAATGAGCCGTCAGACTGCTACAGCATTACC 350  
Db 190 AGAATGAGCCGTCAGACTGCTACAGCATTACC 222  
RESULT 32  
ABL96986  
ID ABL96986 standard; cDNA; 222 BP.  
XX  
XX ABL96986;  
XX  
XX 21-JUN-2002 (first entry)  
XX  
XX Human testicular antigen encoding cDNA SEQ ID NO: 654.  
DE  
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disease; infection; cytostatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200155317-A2.  
XX  
XX 02-AUG-2001.  
PD  
XX

PF 17-JAN-2001; 2001WO-US001329.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216800P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0226683P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.



```
CC in the exemplification of the present invention
XX SQ Sequence 2198 BP; 456 A; 642 C; 579 G; 521 T; 0 U; 0 Other;

Query Match      2.4%; Score 30; DB 6; Length 2198;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 850 GACTGGGTGATGATGAGTCCTGTTTGGC 879
Db 941 GACTGGGTGATGATGAGTCCTGTTTGGC 970

RESULT 34
ADP54673/c
ID ADF54673 standard; DNA; 33 BP.
XX AC
XX ADP54673;
XX 12-FEB-2004 (first entry)
XX DT
XX DE Human SIAH1 PCR primer #2.
XX KW ss; cell proliferation; paternally expressed gene 10; PEG10; cell death;
XX KW cancer; liver cancer; hepatoma; hepatic carcinoma; apoptosis; SIAH1; PCR;
XX KW primer; human.
XX OS Homo sapiens.
XX PN JP2003093066-A.
XX PD 02-APR-2003.
XX PF 21-SEP-2001; 2001JP-00290248.
XX PR 21-SEP-2001; 2001JP-00290248.
XX PA (UYTY ) UNIV TOKYO.
XX PA (ONKO-) ONKO THERAPY SCI KK.
XX DR WPI; 2003-572666/54.
XX KW Promoting or suppressing cell proliferation by increasing or decreasing
XX PT paternally expressed gene 10 (PEG10) protein levels.
XX PS Example 5; SEQ ID NO 10; 25pp; Japanese.
XX CC The invention relates to a method of promoting or suppressing cell
XX CC proliferation by increasing or decreasing paternally expressed gene 10
XX CC (PEG10) protein levels, and suppressing or promoting cell death by
XX CC increasing or decreasing PEG10 protein levels in the cell. The method is
XX CC useful for promoting or suppressing cell proliferation or cell death.
XX CC Preferably, the method is useful for promoting or suppressing
XX CC proliferation or death of cancer cell, preferably liver cancer cell e.g.,
XX CC hepatoma cell. A pharmaceutical composition is useful for treating or
XX CC preventing cell proliferative diseases. The diagnosing method and the
XX CC diagnostic reagent are useful for diagnosing hepatic carcinoma,
XX CC preferably hepatoma. The present sequence is used in the exemplification
XX CC of the invention.
XX SQ Sequence 33 BP; 11 A; 6 C; 7 G; 9 T; 0 U; 0 Other;

Query Match      1.9%; Score 24; DB 10; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1142 GCATCAATGTAACATATTCATGT 1165
Db 33 GCATCAATGTAACATATTCATGT 10

RESULT 35
ABS69096
ID ABS69096 standard; DNA; 226 BP.
XX AC
XX ABS69096;
XX 21-NOV-2002 (first entry)
XX DE Novel murine polynucleotide isolated using gene trap technology #159.
XX KW Mouse; gene-trapped sequence; GTS; functional genomic analysis;
XX KW phage display system; gene chip; temporal gene expression;
XX KW tissue specific gene expression; antisense inhibition; gene targeting;
XX KW development disorder; cell differentiation disorder; aging; cancer;
XX KW autoimmune disease; lupus; inflammatory disorder; skin disorder;
XX KW degenerative disorder; db.
XX OS Mus musculus.
XX PN US2002102543-A1.
XX PD 01-AUG-2002.
XX PF 30-NOV-2000; 2000US-00728445.
XX PR 01-DEC-1999; 99US-0168358P.
XX PA (FRIE/) FRIEDRICH G.
XX PA (ZAMB/) ZAMBROWICZ B.
XX PA (SAND/) SANDS A T.
XX PI Friedrich G, Zambrowicz B, Sands AT;
XX DR WPI; 2002-690598/74.
XX KW Novel murine polynucleotides that individually identify novel genes into
XX PT which a retroviral gene trap vector has integrated, useful in genomic
XX PT analysis and in discovery, development of therapeutic and diagnostic
XX PT agents.
XX PS Claim 1; Page 71; 296pp; English.
XX CC The invention describes an isolated murine polynucleotide (I) comprising
XX CC a contiguous stretch of at least 60 nucleotides of one of 265-677
XX CC nucleotide 891 OMIBANK gene trapped sequences (GTSs) (S), given in the
XX CC specification. The novel genes and cells are useful in functional genomic
XX CC analysis and in the discovery and development of new therapeutic and
XX CC diagnostic agents and methods. (I) is useful for identifying the coding
XX CC regions of the murine genome, to isolate cDNAs, genomic clones, or full-
XX CC length genes/polynucleotides or homologues, heterologues, paralogues, or
XX CC orthologues that are capable of hybridising to one or more of the GTSs
XX CC under stringent conditions. (I) can be incorporated into a phage display
XX CC system that can be used to screen for proteins, or other ligands, that
XX CC are capable of binding an amino acid sequence encoded by an
XX CC oligonucleotide or polynucleotide sequence in at least one of the TS
XX CC sequences. (I) is useful in addressable arrays, such as gene chips, to
XX CC identify and characterise temporal and tissue specific gene expression,
XX CC to identify the gene of interest from many sources and for genetic
XX CC manipulations such as antisense inhibition and gene targeting. Decreasing
XX CC the level of expression of (I) and/or down regulating the activity of
XX CC peptides or proteins encoded by (I) is useful for treating development
XX CC and cell differentiation disorders, aging, cancer, autoimmune disease,
XX CC lupus, inflammatory disorders, skin disorders and degenerative disorders.
XX CC This sequence represents a murine cDNA isolated using gene trap
XX CC technology
XX SQ Sequence 226 BP; 48 A; 69 C; 73 G; 35 T; 0 U; 1 Other;

Query Match      1.8%; Score 23; DB 6; Length 226;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 AAGTGTCCACCATCCAGAGGCT 386
Db 175 AAGTGTCCACCATCCAGAGGCT 197
```

```
RESULT 36
ADF54668
ID ADF54668 standard; DNA; 21 BP.
XX
XX
AC ADF54668;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human SIAH1 reverse transcriptase PCR primer #1.
XX
KW ss; reverse transcriptase; RT-PCR; primer; cell proliferation;
KW paternally expressed gene 10; PEG10; cell death; cancer; liver cancer;
KW hepatoma; hepatic carcinoma; apoptosis; SIAH1; human.
XX
OS Homo sapiens.
XX
PN JP2003093066-A.
XX
PD 02-APR-2003.
XX
PF 21-SEP-2001; 2001JP-00290248.
XX
PR 21-SEP-2001; 2001JP-00290248.
XX
PA (UITY) UNIV TOKYO.
PA (ONKO-) ONKO THERAPY SCI KK.
XX
WPI; 2003-572666/54.
XX
PT Promoting or suppressing cell proliferation by increasing or decreasing
PT paternally expressed gene 10 (PEG10) protein levels.
XX
PS Example 3; SEQ ID NO 5; 25pp; Japanese.
XX
CC The invention relates to a method of promoting or suppressing cell
CC proliferation by increasing or decreasing paternally expressed gene 10
CC (PEG10) protein levels, and suppressing or promoting cell death by
CC increasing or decreasing PEG10 protein levels in the cell. The method is
CC useful for promoting or suppressing cell proliferation or cell death.
CC Preferably, the method is useful for promoting or suppressing
CC proliferation or death of cancer cell, preferably liver cancer cell e.g.,
CC hepatoma cell. A pharmaceutical composition is useful for treating or
CC preventing cell proliferative diseases. The diagnosing method and the
CC diagnostic reagent are useful for diagnosing hepatic carcinoma,
CC preferably hepatoma. The present sequence is used in the exemplification
CC of the invention.
XX
SQ Sequence 21 BP; 6 A; 5 C; 5 G; 5 T; 0 U; 0 Other;
XX
Query Match 1.6%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 412 TCCACAAATGACTTGGCGACT 432
DB 1 TCCACAAATGACTTGGCGACT 21
RESULT 37
ADF54669/c
ID ADF54669 standard; DNA; 21 BP.
XX
XX
AC ADF54669;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human SIAH1 reverse transcriptase PCR primer #2.
XX
KW ss; reverse transcriptase; RT-PCR; primer; cell proliferation;
KW paternally expressed gene 10; PEG10; cell death; cancer; liver cancer;
KW hepatoma; hepatic carcinoma; apoptosis; SIAH1; human.
XX
XX
OS Homo sapiens.
XX
PN JP2003093066-A.
XX
PD 02-APR-2003.
XX
PF 21-SEP-2001; 2001JP-00290248.
XX
PR 21-SEP-2001; 2001JP-00290248.
XX
PA (UITY) UNIV TOKYO.
PA (ONKO-) ONKO THERAPY SCI KK.
XX
WPI; 2003-572666/54.
XX
PT Promoting or suppressing cell proliferation by increasing or decreasing
PT paternally expressed gene 10 (PEG10) protein levels.
XX
PS Example 3; SEQ ID NO 5; 25pp; Japanese.
XX
CC The invention relates to a method of promoting or suppressing cell
CC proliferation by increasing or decreasing paternally expressed gene 10
CC (PEG10) protein levels, and suppressing or promoting cell death by
CC increasing or decreasing PEG10 protein levels in the cell. The method is
CC useful for promoting or suppressing cell proliferation or cell death.
CC Preferably, the method is useful for promoting or suppressing
CC proliferation or death of cancer cell, preferably liver cancer cell e.g.,
CC hepatoma cell. A pharmaceutical composition is useful for treating or
CC preventing cell proliferative diseases. The diagnosing method and the
CC diagnostic reagent are useful for diagnosing hepatic carcinoma,
CC preferably hepatoma. The present sequence is used in the exemplification
CC of the invention.
XX
SQ Sequence 21 BP; 6 A; 5 C; 5 G; 5 T; 0 U; 0 Other;
XX
Query Match 1.6%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 412 TCCACAAATGACTTGGCGACT 432
DB 1 TCCACAAATGACTTGGCGACT 21
RESULT 38
ACH28397
ID ACH28397 standard; cDNA; 442 BP.
XX
XX
AC ACH28397;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human adult ovary cDNA #6777.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
```

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
XX as hybridization probes, as oligomers for PCR, for chromosome and gene  
XX mapping, in the recombinant production of protein, or in generating  
XX antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 15609; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
XX determined by the technique of SBH (sequencing by hybridisation). Also  
XX included is a purified polypeptide comprising a sequence corresponding to  
XX a reading frame of the novel polynucleotide. The nucleic acid sequences  
XX are useful in diagnostics as expressed sequence tags (EST) for  
XX identifying expressed genes or for physical mapping of the human genome,  
XX in forensics, in assessing biodiversity, or in identifying mutations  
XX responsible for genetic disorders and other traits. The nucleotide  
XX sequences are also useful as hybridisation probes, as oligomers for PCR,  
XX for chromosome and gene mapping, in the recombinant production of  
XX protein, or in generating antisense DNA or RNA. The purified polypeptide  
XX is useful for generating antibodies specific for it. The present sequence  
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
XX obtained in electronic format directly from USPTO at  
XX seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
XX Sequence 442 BP; 149 A; 58 C; 92 G; 142 T; 0 U; 1 Other;  
SQ  
Query Match 1.6%; Score 21; DB 9; Length 442;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 165 TTTTAAATATCTTTTAAAT 185  
DB 307 TTTTAAATATCTTTTAAAT 327  
RESULT 39  
ACH28129  
ID ACH28129 standard; cDNA; 477 BP.  
XX  
XX ACH28129;  
XX  
XX 13-OCT-2003 (first entry)  
XX  
XX Human adult ovary cDNA #6509.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
XX genome mapping; biodiversity; genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX US2003073623-A1.  
XX  
XX 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX  
XX (DRNA/) DRMANAC R T.  
XX (LABA/) LABAT I.  
XX (STAC/) STACHE-CRAIN B.  
XX (DICK/) DICKSON M C.  
XX (JONE/) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX WPI; 2003-615964/58.  
XX

XX New polynucleotide sequences obtained from various cDNA libraries, useful  
XX as hybridization probes, as oligomers for PCR, for chromosome and gene  
XX mapping, in the recombinant production of protein, or in generating  
XX antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 15341; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
XX determined by the technique of SBH (sequencing by hybridisation). Also  
XX included is a purified polypeptide comprising a sequence corresponding to  
XX a reading frame of the novel polynucleotide. The nucleic acid sequences  
XX are useful in diagnostics as expressed sequence tags (EST) for  
XX identifying expressed genes or for physical mapping of the human genome,  
XX in forensics, in assessing biodiversity, or in identifying mutations  
XX responsible for genetic disorders and other traits. The nucleotide  
XX sequences are also useful as hybridisation probes, as oligomers for PCR,  
XX for chromosome and gene mapping, in the recombinant production of  
XX protein, or in generating antisense DNA or RNA. The purified polypeptide  
XX is useful for generating antibodies specific for it. The present sequence  
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
XX obtained in electronic format directly from USPTO at  
XX seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
XX Sequence 477 BP; 155 A; 70 C; 98 G; 152 T; 0 U; 2 Other;  
SQ  
Query Match 1.6%; Score 21; DB 9; Length 477;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 165 TTTTAAATATCTTTTAAAT 185  
DB 355 TTTTAAATATCTTTTAAAT 375  
RESULT 40  
ADL14958  
ID ADL14958 standard; DNA; 1094 BP.  
XX  
XX ADL14958;  
XX  
XX 06-MAY-2004 (first entry)  
XX  
XX Human glaucoma-related optineurin (OPTN) exon 12.  
XX  
XX Human; glaucoma; optineurin; OPTN; diagnosis; ds.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX primer\_bind complement(354..380)  
XX /tag= a  
XX exon /note= "Primer SF12"  
XX 501..594  
XX /tag= b  
XX /number= 12  
XX primer\_bind 800..820  
XX /tag= c  
XX /note= "Primer SR12"  
XX  
XX EP1388590-A2.  
XX  
XX 11-FEB-2004.  
XX  
XX 29-JUL-2003; 2003EP-00447201.  
XX  
XX 02-AUG-2002; 2002JP-00226612.  
XX  
XX (SYSM-) SYSMEX CORP.  
XX  
XX Kouchi Y, Masago A, Takahata T;  
XX



XX WPI; 2004-146134/15.  
 XX DR  
 XX Gene assay for predicting future onset of glaucoma, particularly primary  
 PT open angle glaucoma or normal ocular tension glaucoma, comprises  
 PT detecting a mutation of at least one base of the optineurin gene.  
 XX  
 XX Example 1; SEQ ID NO 10; 31pp; English.  
 XX  
 CC The present sequence comprises exon 12 of the glaucoma-associated gene,  
 CC OPTN (optineurin) ADL14949. The invention relates to a gene assay method  
 CC for predicting future onset of primary open angle glaucoma and/or normal  
 CC ocular tension glaucoma. This involves detecting a mutation in the OPTN  
 CC gene coding sequence, specifically a substitution of G for A at position  
 CC 619 and/or a substitution of A for G at position 898 of the OPTN coding  
 CC sequence. The mutation(s) is detected using a nucleic acid amplification  
 CC method using primers specific for the different exons of the coding  
 CC sequence, including primers SF12 ADL14979 and SR12 ADL14980 for exon 12.  
 XX  
 XX Sequence 1094 BP; 316 A; 200 C; 242 G; 336 T; 0 U; 0 Other;  
 SQ  
 Query Match 1.6%; Score 21; DB 12; Length 1094;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 165 TTTTAAATATCTTTTAAAT 185  
 DB 478 TTTTAAATATCTTTTAAAT 498  
 RESULT 41  
 ID ADL14957  
 ID ADL14957 standard; DNA; 1150 BP.  
 XX  
 AC ADL14957;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Human glaucoma-related optineurin (OPTN) exon 11.  
 XX  
 KW Human; glaucoma; optineurin; OPTN; diagnosis; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT primer\_bind /\*tag= a  
 FT complement (325..344)  
 FT exon /\*note= "Primer SF11"  
 FT 501..650  
 FT /\*tag= b  
 FT primer\_bind /number= 11  
 FT 834..853  
 FT /\*tag= c  
 FT /\*note= "Primer SR11"  
 XX  
 DN EPI388590-A2.  
 XX  
 XX 11-FEB-2004.  
 XX  
 XX 29-JUL-2003; 2003EP-00447201.  
 XX  
 XX 02-AUG-2002; 2002JP-00226612.  
 XX  
 XX (SYSM-) SYSMEX CORP.  
 XX  
 XX Kouchi Y, Masago A, Takahata T;  
 XX  
 XX WPI; 2004-146134/15.  
 XX  
 XX Gene assay for predicting future onset of glaucoma, particularly primary  
 PT open angle glaucoma or normal ocular tension glaucoma, comprises  
 PT detecting a mutation of at least one base of the optineurin gene.  
 XX

PS Example 1; SEQ ID NO 9; 31pp; English.  
 XX  
 CC The present sequence comprises exon 11 of the glaucoma-associated gene,  
 CC OPTN (optineurin) ADL14949. The invention relates to a gene assay method  
 CC for predicting future onset of primary open angle glaucoma and/or normal  
 CC ocular tension glaucoma. This involves detecting a mutation in the OPTN  
 CC gene coding sequence, specifically a substitution of G for A at position  
 CC 619 and/or a substitution of A for G at position 898 of the OPTN coding  
 CC sequence. The mutation(s) is detected using a nucleic acid amplification  
 CC method using primers specific for the different exons of the coding  
 CC sequence, including primers SF11 ADL14977 and SR11 ADL14977 for exon 11.  
 XX  
 XX Sequence 1150 BP; 381 A; 172 C; 229 G; 368 T; 0 U; 0 Other;  
 SQ  
 Query Match 1.6%; Score 21; DB 12; Length 1150;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 165 TTTTAAATATCTTTTAAAT 185  
 DB 1006 TTTTAAATATCTTTTAAAT 1026  
 RESULT 42  
 ID ADE13891  
 ID ADE13891 standard; DNA; 46951 BP.  
 XX  
 AC ADE13891;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human optineurin gene.  
 XX  
 KW Human; optineurin; ds; gene; ophthalmological;  
 KW single nucleotide polymorphism; SNP; glaucoma;  
 KW progressive ocular hypertensive disorder; glaucoma related disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT variation replace(391,G)  
 FT /\*tag= a  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(691,G)  
 FT /\*tag= b  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(709,G)  
 FT /\*tag= c  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(887,A)  
 FT /\*tag= d  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(894,T)  
 FT /\*tag= e  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(987,C)  
 FT /\*tag= f  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(1112,C)  
 FT /\*tag= g  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(1505,CC)  
 FT /\*tag= h  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(1606,A)  
 FT /\*tag= i  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(2405,T)  
 FT /\*tag= j  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(2606,G)  
 FT /\*tag= k  
 FT standard\_name= "Single nucleotide polymorphism"  
 FT

FT variation replace(3313,A)  
 FT /tag= 1  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(3555,TT)  
 FT /tag= m  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(3625,G)  
 FT /tag= n  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(3629,C)  
 FT /tag= o  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(3882,TT)  
 FT /tag= p  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(3988,T)  
 FT /tag= q  
 FT /standard\_name= "Single nucleotide polymorphism"  
 FT variation replace(4452,A)  
 FT /tag= r  
 FT /standard\_name= "Single nucleotide polymorphism"  
 XX US2003190617-A1.  
 PN  
 XX  
 XX 09-OCT-2003.  
 XX  
 XX 06-MAR-2002; 2002US-00091281.  
 XX  
 XX 06-MAR-2002; 2002US-00091281.  
 XX  
 XX (SIEE/) SI E.  
 PA (RAYM/) RAYMOND V.  
 PA (MORI/) MORISSETTE J.  
 XX  
 XX Raymond V, Morissette J, Si E;  
 XX WPI; 2003-864168/80.  
 DR  
 XX  
 XX New nucleic acid sequences of the optineurin gene are useful to detect  
 FT polymorphisms particularly single nucleotide polymorphisms in the  
 FT optineurin promoter to diagnose, prognose and treat glaucoma and related  
 FT disorders.  
 XX  
 XX Disclosure; SEQ ID NO 2; 159pp; English.  
 PS  
 CC The invention relates to an isolated nucleic acid (N1) comprising at  
 CC least 20 but not more than 1500 consecutive nucleotides of the optineurin  
 CC promoter appearing as ADB1390. Also included are the optineurin promoter  
 CC operably linked to a heterologous nucleic acid, a nucleic acid capable of  
 CC detecting a single nucleotide polymorphism (SNP) in the optineurin  
 CC promoter, a host cell comprising the promoter operably linked to a  
 CC heterologous sequence, diagnosing or prognosing glaucoma in a sample  
 CC obtained from a cell or bodily fluid (comprising glaucoma in a sample  
 CC in a promoter region of the optineurin gene, associated with a glaucoma  
 CC phenotype), detecting a SNP sequence variation in a sample containing  
 CC DNA, detecting the presence of an optineurin promoter sequence variation  
 CC in a sample containing DNA, determining the presence or increased  
 CC susceptibility to glaucoma or to a progressive ocular hypertensive  
 CC disorder resulting in loss of visual field in a patient (or the severity  
 CC or progression of glaucoma in a patient, comprising providing  
 CC amplification reaction primers that direct amplification of a selected  
 CC nucleic acid region containing the variation within the optineurin  
 CC promoter and amplifying the DNA) and detecting a polymorphism (comprising  
 CC obtaining a sample containing human genomic DNA, providing a nucleic acid  
 CC capable of detecting a SNP located within an optineurin promoter, and  
 CC detecting the polymorphism). The invention is used to diagnose and  
 CC prognose glaucoma and also to treat glaucoma related disorders. The  
 CC present sequence is the optineurin gene.  
 XX  
 SQ Sequence 46951 BP; 12703 A; 10108 C; 10051 G; 14089 T; 0 U; 0 Other;

Query Match 1.6%; Score 21; DB 10; Length 46951;  
 Best Local Similarity 100.0%; Pred. No. 13;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 165 TTTTAAATATCTTTTAAAT 185  
 |||||  
 Db 30767 TTTTAAATATCTTTTAAAT 30787  
 |||||  
 RESULT 43  
 AAC67288  
 ID AAC67288 standard; DNA; 28 BP.  
 XX  
 XX AAC67288;  
 XX  
 XX 09-APR-2001 (first entry)  
 XX  
 XX Human Siah-1alpha PCR primer SEQ ID NO: 42.  
 XX  
 XX Human; protein degradation; siah-mediated degradation protein; SMWP;  
 KW SCP-complex protein; SCP; siah-1alpha; siah-1 interacting protein; SIP;  
 KW Skp1-associated F-box protein; SAF-1; SAF-2; SAD; cancer; cell division;  
 KW Skp1-associated destruction-box protein; inflammatory disease;  
 KW PCR primer; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200077207-A2.  
 PN  
 XX 21-DEC-2000.  
 PD  
 XX  
 XX 09-JUN-2000; 2000WO-US015873.  
 PF  
 XX 11-JUN-1999; 99US-00330517.  
 PR  
 XX (BURN-) BURNHAM INST.  
 PA  
 XX Reed JC, Matsuzawa S;  
 PI WPI; 2001-071273/08.  
 DR  
 XX Siah-Mediated Degradation Protein, useful for drug screening, for  
 FT therapeutic applications and for functional genomics.  
 PT  
 XX Example 2; Page 58; 121pp; English.  
 PS  
 XX The present invention provides the protein and coding sequences of  
 CC several siah-mediated degradation proteins and SCP-complex proteins.  
 CC These are designated Siah-1alpha, Siah-1 interacting protein (SIP), which  
 CC encodes two proteins due to alternative splicing (SIP-L and SIP-S), Skp1-  
 CC associated F-box protein-alpha and beta and -2 (SAF-1alpha, SAF-1beta  
 CC and SAF-2) and Skp1-associated destruction-box protein (SAD). The  
 CC proteins and their coding sequences are useful in the diagnosis and  
 CC treatment of cancers, disorders where too little cell division occurs  
 CC such as bone marrow aplasia, immunodeficiencies and inflammatory  
 CC diseases including sepsis, fibrosis, arthritis and graft versus host  
 CC disease  
 XX  
 SQ Sequence 28 BP; 9 A; 4 C; 8 G; 7 T; 0 U; 0 Other;  
 Query Match 1.6%; Score 20; DB 4; Length 28;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 236 GGACTTATGGCATGTAAACA 255  
 |||||  
 Db 9 GGACTTATGGCATGTAAACA 28  
 |||||  
 RESULT 44  
 ACA55799  
 ID ACA55799 standard; cDNA; 257 BP.  
 XX  
 XX ACA55799;  
 AC  
 XX

DT 06-JUN-2003 (first entry)  
XX Human signalling pathway polynucleotide probe SEQ ID NO 397.  
DE  
XX  
XX Human; probe; ss; array element; Parkinson's disease;  
KW signalling pathway polynucleotide; cancer; adenocarcinoma; leukaemia;  
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.  
XX  
OS Homo sapiens.  
XX  
XX US6500938-B1.  
XX  
XX 31-DEC-2002.  
XX  
XX 30-JAN-1998; 98US-00016434.  
XX  
XX 30-JAN-1998; 98US-00016434.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Au-Young J, Seilhamer JJ;  
XX WPI; 2003-352189/33.  
XX  
XX Combination of polynucleotide probes, useful as array elements in a  
PT microarray for monitoring the expression of a number of target  
PT polynucleotides.  
XX  
XX Claim 1; SEQ ID NO 397; 65pp; English.  
XX  
XX The invention relates to a combination which, comprises a number of  
CC polynucleotide probes comprising a sequence selected from one of the 1490  
CC sequences mentioned in the specification. The combination is useful as an  
CC array element in a microarray for monitoring the expression of a number  
CC of target polynucleotides. The microarray is particularly useful in the  
CC diagnosis and treatment of cancer and immunopathology and neuropathology.  
CC The microarray is useful in diagnostics and treatment regimens, drug  
CC discovery and development, toxicological and carcinogenicity studies,  
CC forensics and pharmacogenomics. The microarray is also useful for  
CC monitoring progression of diseases and for developing sophisticated  
CC profiles for the effects of currently available therapeutic drugs. The  
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs  
CC and genomic fragments and in research and diagnostic applications. The  
CC array can detect changes in expression in a large number of genes coding  
CC for different signaling pathway populations which can be used to diagnose  
CC various diseases including cancer e.g. adenocarcinoma and leukaemia.  
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease  
CC and Parkinson's disease. The present sequence represents a polynucleotide  
CC probe of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=06500938B1  
XX  
XX Sequence 257 BP; 59 A; 75 C; 63 G; 60 T; 0 U; 0 Other;  
SQ  
Query Match 1.6%; Score 20; DB 10; Length 257;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 850 GACTGGGTGATGATGCAGTC 869  
Db 222 GACTGGGTGATGATGCAGTC 241  
RESULT 45  
ADI55595  
ID ADI55595 standard; DNA; 257 BP.  
XX  
XX AC ADI55595;  
XX  
XX 22-APR-2004 (first entry)  
XX  
XX DE Human polynucleotide probe #397.

XX Human; probe; ss; receptor-like polypeptide; transducing polypeptide;  
KW effector-like polypeptide; cancer; immunopathology; neuropathology;  
KW drug development; toxicology; carcinogenicity;  
KW signalling pathway polypeptide; adrenal gland; bladder; bone;  
KW bone marrow; brain; breast; cervix; tumour; immunopathology; AIDS;  
KW diabetes; pancreatitis; osteoporosis; ulcerative colitis; neuropathology;  
KW dementia; amnesia; epilepsy; Alzheimer's disease; depression.  
XX  
OS Homo sapiens.  
XX  
XX US2004010136-A1.  
XX  
XX 15-JAN-2004.  
XX  
XX 26-NOV-2002; 2002US-00305720.  
XX  
XX 30-JAN-1998; 98US-00016434.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Au-Young J, Seilhamer JJ;  
XX WPI; 2004-090520/09.  
XX  
XX New composition comprising polynucleotide probes, useful as array  
PT elements in a microarray for monitoring the expression of target  
PT polynucleotides or purifying a subpopulation of mRNAs, cDNA, or genomic  
PT fragments.  
XX  
XX Claim 6; SEQ ID NO 397; 73pp; English.  
XX  
XX The invention relates to a composition of polynucleotide probes  
CC comprising first polynucleotide probes comprising at least a portion of a  
CC gene encoding a receptor-like polypeptide, second polynucleotide probes  
CC comprising at least a portion of a gene encoding a transducing  
CC polypeptide and third polynucleotide probes comprising at least a portion  
CC of a gene encoding an effector-like polypeptide. The probes of the  
CC composition are useful as array elements in a microarray for monitoring  
CC the expression of target polynucleotides. The microarray is useful in the  
CC diagnosis and treatment of cancer, an immunopathology or a  
CC neuropathology. It can also be used for drug discovery and development,  
CC toxicological and carcinogenicity studies, forensics or pharmacogenomics.  
CC Microarrays can also be used for monitoring the progression of diseases  
CC that may be associated with the altered expression of signalling pathway  
CC polypeptides. The composition can also be used to purify a subpopulation  
CC of mRNAs, cDNAs, or genomic fragments in a sample. The expression profile  
CC is also useful for the diagnosis and treatment of cancer, e.g. cancers of  
CC the adrenal gland, bladder, bone, bone marrow, brain, breast or cervix,  
CC an immunopathology, e.g. AIDS, diabetes, pancreatitis, osteoporosis or  
CC ulcerative colitis, or a neuropathology, e.g. dementia, amnesia,  
CC epilepsy, Alzheimer's disease or depression. This sequence represents a  
CC human polynucleotide probe of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html.  
XX  
XX Sequence 257 BP; 59 A; 75 C; 63 G; 60 T; 0 U; 0 Other;  
SQ  
Query Match 1.6%; Score 20; DB 12; Length 257;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 850 GACTGGGTGATGATGCAGTC 869  
Db 222 GACTGGGTGATGATGCAGTC 241  
RESULT 46  
ABL08593  
ID ABL08593 standard; cDNA; 834 BP.  
XX  
XX AC ABL08593;

XX DT 26-MAR-2002 (first entry)  
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 20261.  
XX KW Drosophila; developmental biology; cell signalling; insecticide;  
XX OS pharmaceutical; gene; ss.  
XX PN Drosophila melanogaster.  
XX PD WO200171042-A2.  
XX PF 27-SEP-2001.  
XX PS 23-MAR-2001; 2001WO-US009231.  
XX PR 23-MAR-2000; 2000US-0191637P.  
XX PA 11-JUL-2000; 2000US-00614150.  
XX PI (PEKE ) PE CORP NY.  
XX PT Venter JC, Adams M, Li PWD, Myers EW;  
XX DR WPI; 2001-656860/75.  
XX DR P-PSDB; ABB64490.  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX PT interactions.  
XX PS Claim 1; SEQ ID NO 20261; 21pp + Sequence Listing; English.  
XX CC The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is  
XX CC useful in developmental biology and in elucidating cell signalling and  
XX CC cell-cell interactions in higher eukaryotes for the development of  
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention  
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
XX CC ABB72072). The sequence data for this patent did not form part of the  
XX CC printed specification, but was obtained in electronic format directly  
XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 834 BP; 255 A; 180 C; 165 G; 234 T; 0 U; 0 Other;  
  
Query Match 1.6%; Score 20; DB 4; Length 834;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 247 ATGTAACATTATTATATAA 266  
DB 49 ATGTAACATTATTATATAA 68  
  
RESULT 47  
ADR25256  
ID ADR25256 standard; DNA; 975 BP.  
AC ADR25256;  
XX 21-OCT-2004 (first entry)  
DT Breast cancer prognosis marker #1117.  
DE Breast cancer; prognosis; gene expression; diagnosis.  
KW ds; breast cancer; prognosis; gene expression; diagnosis.  
XX Homo sapiens.  
XX WO2004065545-A2.  
PN 05-AUG-2004.  
PD 15-JAN-2004; 2004WO-US001100.

XX 15-JAN-2003; 2003US-00342887.  
XX (ROSE-) ROSETTA INPHARMATICS LLC.  
XX PA (NECA-) NETHERLANDS CANCER INST.  
XX Van't Veer LJ, He Y;  
XX WPI; 2004-593473/57.  
XX Classifying a breast cancer patient according to prognosis comprises  
XX PT determining the similarity between the level of expression of each of  
XX PT five genes in a cell sample taken from patient, to control levels.  
XX PS Disclosure; SEQ ID NO 1117; 226pp; English.  
XX CC The invention relates to a method of classifying a breast cancer patient  
XX CC according to prognosis by determining the similarity between the level of  
XX CC expression of each of five genes for which markers are listed in the  
XX CC specification, in a cell sample taken from the breast cancer patient, to  
XX CC control levels of expression for each respective five genes to obtain a  
XX CC patient similarity value. The methods are useful for classifying a breast  
XX CC cancer patient according to prognosis. Kits and computer program products  
XX CC are useful for data analysis using the diagnostic, prognostic and  
XX CC statistical methods of the invention. This sequence corresponds to a  
XX CC marker used in the method of the invention.  
XX SQ Sequence 975 BP; 188 A; 316 C; 271 G; 200 T; 0 U; 0 Other;  
  
Query Match 1.6%; Score 20; DB 13; Length 975;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 850 GACTGGTGATGATGCAGTC 869  
DB 649 GACTGGTGATGATGCAGTC 668  
  
RESULT 48  
ACC50292  
ID ACC50292 standard; cDNA; 2240 BP.  
XX AC ACC50292;  
XX 12-JUN-2003 (first entry)  
DT Breast cancer associated cDNA sequence SEQ ID NO:424.  
DE Human; breast cancer; cytostatic; gene therapy; gene; ss.  
XX Homo sapiens.  
XX WO2003004989-A2.  
XX 16-JAN-2003.  
XX 21-JUN-2002; 2002WO-US019669.  
XX 21-JUN-2001; 2001US-0299887P.  
XX 27-JUN-2001; 2001US-0301572P.  
XX 18-JUL-2001; 2001US-0306501P.  
XX 25-SEP-2001; 2001US-0325002P.  
XX 05-MAR-2002; 2002US-0362585P.  
XX 14-MAY-2002; 2002US-0380391P.  
XX (MILL-) MILLENIUM PHARM INC.  
XX Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;  
XX PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;  
XX PI Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;  
XX WPI; 2003-210381/20.  
XX P-PSDB; ABR47593.

XX Breast cancer diagnosis or treatment by comparing the level of expression  
PT of a marker in a patient sample with that in the control non-breast  
PT cancer sample.  
XX  
XX Claim 1; SEQ ID NO 424; 128pp; English.  
XX  
CC The present invention describes a method for assessing whether a patient  
CC is afflicted with breast cancer. The method comprises comparing the level  
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and  
CC ABR47386 to ABR47632) in a patient sample and the normal level of  
CC expression of the marker in a control non-breast cancer sample, where a  
CC significant increase in the level of expression of the marker in the  
CC patient sample and the normal level is an indication that the patient is  
CC afflicted with breast cancer. The breast cancer associated sequences from  
CC the present invention have cytostatic activities and can be used in gene  
CC therapy. The method is useful for diagnosing and treating breast cancer.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 2240 BP; 429 A; 676 C; 637 G; 498 T; 0 U; 0 Other;  
SQ  
Query Match 1.6%; Score 20; DB 8; Length 2240;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 850 GACTGGGTGATGATGCAGTC 869  
DB 1175 GACTGGGTGATGATGCAGTC 1194  
RESULT 49  
ADP13401  
ID ADP13401 standard; DNA; 2240 BP.  
XX  
XX ADP13401;  
XX  
XX 26-AUG-2004 (first entry)  
XX  
XX Renal cell carcinoma differentially expressed gene #137.  
XX  
XX ds; diagnosis; non-blood disease; solid tumor; gene expression;  
KW peripheral blood mononuclear cell; renal cell carcinoma; prostate cancer;  
KW head/neck cancer; differential expression.  
XX  
XX Homo sapiens.  
XX  
XX WO2004048933-A2.  
XX  
XX 10-JUN-2004.  
XX  
XX 21-NOV-2003; 2003WO-US037481.  
XX  
XX 21-NOV-2002; 2002US-0427982P.  
XX  
XX 03-APR-2003; 2003US-0459782P.  
XX  
XX (AMEP ) WYETH.  
XX (TWIN/) TWINE N C.  
XX (BUREC/) BURCZYNSKI M E.  
XX (TREF/) TREPICCHIO W L.  
XX (DORN/) DORNER A.  
XX (STOV/) STOVER J A.  
XX (SLON/) SLONI D K.  
XX  
XX Twine NC, Burczynski ME, Trepicchio WL, Dornier A, Stover JA;  
PI Sloni DK;  
XX  
XX WPI; 2004-460799/43.  
XX  
XX Diagnosing non-blood disease such as solid tumor, involves comparing  
PT differential expression profile of specific genes in peripheral blood  
PT sample of subject with reference expression profile of specific genes.

XX Disclosure; SEQ ID NO 137; 350pp; English.  
XX  
XX The invention relate to a method of diagnosing (M1) non-blood disease  
CC such as solid tumor by providing peripheral blood sample of human having  
CC non-blood disease, and comparing an expression profile of specific genes  
CC in the peripheral blood sample to reference expression profile of the  
CC genes, where each of the genes is differentially expressed in peripheral  
CC blood mononuclear cells (PBMCs) of patients having the disease as  
CC compared to PBMCs of normal humans. The method is useful for diagnosing  
CC non-blood disease such as solid tumor. The solid tumor is chosen from  
CC renal cell carcinoma (RCC), prostate cancer and head/neck cancer. The  
CC peripheral blood sample comprises enriched PBMCs. The peripheral blood  
CC sample is a whole blood sample (claimed). (M1) is useful for identifying  
CC genes that are differentially expressed in peripheral blood samples  
CC isolated at different stages of progression, development or treatment of  
CC RCC and/or other solid tumors. This sequence corresponds to a gene that  
CC is differentially expressed and detected by the method of the invention.  
CC (Note: this sequence is not given as part of the printed specification  
CC but was obtained from WIPO in electronic format at  
CC ftp.wipo.int/pub/published\_pct\_sequences).  
XX  
XX Sequence 2240 BP; 429 A; 676 C; 637 G; 498 T; 0 U; 0 Other;  
SQ  
Query Match 1.6%; Score 20; DB 12; Length 2240;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 850 GACTGGGTGATGATGCAGTC 869  
DB 1175 GACTGGGTGATGATGCAGTC 1194  
RESULT 50  
ADP14122  
ID ADP14122 standard; DNA; 2240 BP.  
XX  
XX ADP14122;  
XX  
XX 21-OCT-2004 (first entry)  
XX  
XX Human NF-kappaB pathway-associated gene SeqID123.  
XX  
XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;  
KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;  
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;  
KW immunosuppressive; vulnery; gene therapy; immune disorder;  
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;  
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;  
KW hyper-igm syndrome; hypohidrotic ectodermal dysplasia;  
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;  
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;  
KW viral replication; host cell survival; evasion of immune response;  
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;  
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;  
KW autoimmune disorder; hyper immune activity;  
KW aberrant acute phase response; hypertransplant rejection;  
KW necrotic lesion; wound; organ transduction; proliferating disorder; cancer;  
KW HIV propagation; gene; ds; human.  
XX  
XX Homo sapiens.  
XX  
XX WO2004065577-A2.  
XX  
XX 05-AUG-2004.  
XX  
XX 13-JAN-2004; 2004WO-US000798.  
XX  
XX 14-JAN-2003; 2003US-0440068P.  
XX  
XX 12-MAY-2003; 2003US-0469757P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.



*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005: Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 00:27:47 ; Search time 4597 Seconds  
(without alignments)  
10549.027 Million cell updates/sec

Title: US-10-679-246-1  
Perfect score: 1274  
Sequence: 1 tttctttagtctttagt.....ctctttagttagtggaagc 1274

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 15

Total number of hits satisfying chosen parameters: 102688

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsa1:\*

9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1062	83.4	1221	7	U70056 U70056 Soar
2	958	75.2	1360	3	CR607323 full-length
3	958	75.2	1857	3	CR614925 full-length
4	894	70.2	1287	3	CR598465 full-length
5	849	66.6	849	9	AY402255 Homo sapi
6	685	53.8	1576	3	CR593197 full-length
7	666	52.3	714	5	BQ447177 UI-H-EUI-
8	666	52.3	769	5	BQ447177 UI-H-EUI-
9	645	50.6	849	9	AY402256 Pan trogl
10	642	50.4	812	1	AUI119916 AUI19916
11	637	50.0	859	5	EX366404 EX366404
12	633	49.7	693	5	BH662334
13	624	49.0	824	4	BG619351 C184e12.z
14	622	48.8	622	7	CV030563 9727 Full
15	615	48.3	649	2	BF446537 7g88f02.x
16	611	48.0	779	1	A1936347 wo81b11.x
17	590	46.3	735	4	B1560892 603254068
18	561	44.0	778	1	A1971258
19	525	41.2	980	6	CD251717 AGENCOURT
20	502	39.4	586	7	CK822873 i119b01.y
21	499	39.2	734	1	AUI138740 AUI138740
22	498	39.1	852	5	EX463976 EX463976
23	497	39.0	969	5	EX390856 EX390856
24	494	38.8	637	1	AI681134 tx44c10.x

25	494	38.8	757	1	AUI36155
26	487	38.2	1157	5	BQ67628 AGENCOURT
27	486	38.1	562	1	AUI46134
28	486	38.1	767	4	BG435333 602508085
29	485	38.1	647	2	BE502849 h280b11.x
30	485	38.1	976	5	BUI191771 AGENCOURT
31	477	37.4	890	7	CF551876 AGENCOURT
32	475	37.3	515	1	AI803300
33	475	37.3	785	5	BE567663 601340316
34	475	37.3	785	5	BE567663 601340316
35	474	37.2	647	2	BE503207 h284b07.x
36	474	37.2	800	4	BI836179
37	473	37.1	842	5	EX459114
38	473	37.1	894	5	BP227842
39	470	36.9	583	5	BP227842
40	462	36.3	957	1	BX372197
41	459	36.0	513	1	AA917325
42	457	35.9	658	4	BG537816
43	456	35.8	608	1	AI698102
44	452	35.5	487	2	BF839462
45	440	34.5	561	1	AA418482
46	430	33.8	1058	5	BX356371
47	429	33.7	882	5	BX340272
48	428	33.6	578	5	BP332871
49	419	32.9	853	4	BG436513
50	411	32.3	913	5	BQ431458
51	404	31.7	445	1	AI284188
52	401	31.5	438	1	AA971734
53	400	31.4	550	5	BQ365979
54	397	31.2	567	5	BQ086148
55	392	30.8	497	1	AA258553
56	389	30.5	652	2	AW274939
57	387	30.4	419	1	AA054272
58	387	30.4	815	2	BF033914
59	382	30.0	578	5	BP243635
60	376	29.5	376	4	BM699360
61	371	29.1	377	1	AA455956
62	370	29.0	736	6	CA427778
63	369	29.0	797	2	BF102716
64	367	28.8	864	4	BG489183
65	365	28.6	1025	5	BX366403
66	363	28.5	1022	5	BX380468
67	360	28.3	697	2	AW390258
68	357	28.0	499	1	AA054304
69	348	27.3	715	5	BX459079
70	346	27.2	760	1	AUI136857
71	344	27.0	442	1	AA075461
72	343	26.9	455	7	W79008
73	332	26.1	375	7	H02936
74	328	25.7	418	7	H72152
75	326	25.6	952	5	BX357983
76	322	25.3	902	4	BI763325
77	319	25.0	872	5	BX448317
78	311	24.4	490	1	AA447531
79	311	24.4	617	5	BG657753
80	310	24.3	492	1	AA022943
81	308	24.2	924	5	BUS70869
82	308	24.2	1010	4	BM800208
83	304	23.9	834	6	CD106578
84	303	23.8	892	5	BUS69099
85	300	23.5	895	2	BF769906
86	298	23.4	453	1	AI304442
87	293	23.0	582	5	BP330742
88	293	23.0	844	2	BF185881
89	293	23.0	921	5	BX355841
90	292	22.9	292	7	DS9180
91	292	22.9	435	2	AA813663
92	289	22.7	478	1	AA813663
93	288	22.6	850	4	BG500464
94	286	22.4	880	7	CF593249
95	285	22.4	994	5	BX339645
96	284	22.3	740	6	CD250494
97	283	22.2	836	6	CB994051







C 536	20	1.6	618	9	AG230190	Lotus cor	609	20	1.6	806	7	CR417945	CR417945
C 537	20	1.6	622	1	AL674749	AL674749	C 610	20	1.6	809	7	CG304555	CG304555
C 538	20	1.6	625	1	AV841171	AV841171	611	20	1.6	818	9	CR582937	CR582937
C 539	20	1.6	632	4	BM790493	K-EST0070	C 612	20	1.6	825	8	BZ255361	BZ255361
540	20	1.6	641	4	BI168259	R08857.5	613	20	1.6	836	9	CC826591	CC826591
541	20	1.6	642	5	BM346293	BM346293	614	20	1.6	846	9	CC485382	CC485382
542	20	1.6	647	2	BE581998	K457809.Y	615	20	1.6	847	5	BQ220310	BQ220310
543	20	1.6	647	4	EJ697375	BJ697375	C 616	20	1.6	848	9	CC654797	CC654797
544	20	1.6	648	5	BW255679	BW255679	C 617	20	1.6	856	9	CL839696	CL839696
545	20	1.6	649	8	AZ615790	1M0445L13	618	20	1.6	858	9	CG9694501	CG9694501
C 546	20	1.6	650	9	CE033761	tigr-gss-	C 619	20	1.6	859	2	BF272342	BF272342
547	20	1.6	657	2	BE580308	K319f10.Y	C 620	20	1.6	868	5	EX736384	EX736384
548	20	1.6	658	9	CC619652	OCWC192TH	C 621	20	1.6	873	9	CL250563	CL250563
549	20	1.6	659	5	BW329467	BW329467	C 622	20	1.6	874	9	CG433160	CG433160
C 550	20	1.6	659	9	CG871446	ZMMBBc027	C 623	20	1.6	882	9	CG739325	CG739325
551	20	1.6	664	8	BZ955499	CH240.961	C 624	20	1.6	885	5	BU164429	BU164429
552	20	1.6	666	7	CK023635	AGENCOURT	C 625	20	1.6	886	9	CG235556	CG235556
553	20	1.6	676	8	AZ574928	334PvD02	C 626	20	1.6	890	5	CG331038	CG331038
554	20	1.6	676	9	CC834134	ZMMBBc018	C 627	20	1.6	902	9	CG874395	CG874395
C 555	20	1.6	680	5	BW186062	BW186062	C 628	20	1.6	914	5	EX772276	EX772276
556	20	1.6	680	7	CR435389	CR435389	C 629	20	1.6	915	9	CG331038	CG331038
557	20	1.6	682	4	BJ689168	BJ689168	C 630	20	1.6	928	9	CL991827	CL991827
558	20	1.6	689	4	BM551946	BM551946	C 631	20	1.6	931	9	CL317081	CL317081
559	20	1.6	692	4	BI909621	AGENCOURT	C 632	20	1.6	939	9	CL447543	CL447543
C 560	20	1.6	693	1	AV866684	AV866684	C 633	20	1.6	952	9	CG90405	CG90405
561	20	1.6	700	9	CG917869	ZMMBBc038	C 634	20	1.6	953	9	CL457197	CL457197
562	20	1.6	706	9	CC832038	ZMMBBb018	635	20	1.6	954	8	CC196243	CC196243
563	20	1.6	707	1	AA697948	HL03517.5	C 636	20	1.6	957	9	CG420847	CG420847
564	20	1.6	708	9	CR330078	CR330078	C 637	20	1.6	960	9	CG103746	CG103746
565	20	1.6	709	9	CC877682	ZMMBBb021	C 638	20	1.6	960	9	CG908751	CG908751
566	20	1.6	710	8	BZ396230	BZ396230	639	20	1.6	981	9	CG260327	CG260327
567	20	1.6	713	9	AG556257	Mus muscu	640	20	1.6	981	9	CG331038	CG331038
C 568	20	1.6	714	6	CD781337	EST652698	C 641	20	1.6	992	8	CC314927	CC314927
569	20	1.6	718	5	BW383540	BW383540	C 642	20	1.6	993	9	CG900859	CG900859
570	20	1.6	718	7	CN925584	000512AEN	C 643	20	1.6	995	9	CC973749	CC973749
571	20	1.6	721	7	CN925584	000512AEN	C 644	20	1.6	996	9	CL235504	CL235504
C 572	20	1.6	727	9	CE448504	tigr-gss-	645	20	1.6	1002	5	BX355762	BX355762
C 573	20	1.6	727	9	CE448504	tigr-gss-	C 646	20	1.6	1008	9	CG876520	CG876520
C 574	20	1.6	730	5	BW417324	BW417324	C 647	20	1.6	1012	8	BZ746420	BZ746420
C 575	20	1.6	730	9	CL155719	104.342.1	C 648	20	1.6	1012	8	CL292438	CL292438
C 576	20	1.6	732	9	CG824172	ZMMBBb016	649	20	1.6	1016	3	CR683141	CR683141
C 577	20	1.6	734	5	BW395908	BW395908	C 650	20	1.6	1016	5	BX370682	BX370682
C 578	20	1.6	735	5	CG798015	CG798015	C 651	20	1.6	1032	9	CL446232	CL446232
C 579	20	1.6	736	5	BP714500	BP714500	C 652	20	1.6	1058	9	CL082094	CL082094
580	20	1.6	736	7	CK978640	CK978640	C 653	20	1.6	1078	9	CNS06890	CNS06890
581	20	1.6	738	9	CC750658	ZMMBBb013	654	20	1.6	1130	1	AL561412	AL561412
C 582	20	1.6	739	8	BH327596	CH230-174	C 655	20	1.6	1133	9	CG878461	CG878461
C 583	20	1.6	739	9	CL287086	ZMMBBb063	C 656	20	1.6	1241	9	CL648737	CL648737
C 584	20	1.6	741	5	EX696845	EX696845	C 657	20	1.6	1241	9	CL647711	CL647711
585	20	1.6	742	9	CC828120	ZMMBBb017	C 658	20	1.6	1243	4	BG430724	BG430724
C 586	20	1.6	747	5	BW330540	BW330540	659	20	1.6	1243	3	CR610024	CR610024
C 587	20	1.6	748	5	BW400395	BW400395	660	20	1.6	1337	5	BP509566	BP509566
C 588	20	1.6	753	5	BW272269	BW272269	661	20	1.5	169	7	CK907022	CK907022
589	20	1.6	755	9	CL237651	CL237651	C 662	20	1.5	175	8	AZ034720	AZ034720
C 590	20	1.6	760	8	CC338732	OGIAS17TV	C 663	20	1.5	189	7	CNS73991	CNS73991
C 591	20	1.6	761	9	CL381974	RPC144.32	664	20	1.5	190	4	BI814537	BI814537
C 592	20	1.6	766	6	CD559426	AGENCOURT	665	20	1.5	200	4	BM446610	BM446610
593	20	1.6	767	5	BU144068	603229376	666	20	1.5	241	1	AV143692	AV143692
C 594	20	1.6	767	7	CF217266	AGENCOURT	C 667	20	1.5	241	4	BG734977	BG734977
C 595	20	1.6	773	5	BW129504	BW129504	C 668	20	1.5	243	7	CV246887	CV246887
C 596	20	1.6	773	5	BW387717	BW387717	C 669	20	1.5	253	9	CE033070	CE033070
C 597	20	1.6	774	5	BW079214	BW079214	670	20	1.5	277	6	CA302129	CA302129
C 598	20	1.6	774	5	BW397735	BW397735	671	20	1.5	301	9	CE516406	CE516406
C 599	20	1.6	776	8	BZ969506	PUI109TB	C 672	20	1.5	316	2	BB222262	BB222262
C 600	20	1.6	778	5	BW010130	BW010130	673	20	1.5	316	9	CE785938	CE785938
601	20	1.6	781	9	CG883553	CG883553	674	20	1.5	323	9	CE641681	CE641681
C 602	20	1.6	785	4	BG217318	RST37023	675	20	1.5	333	6	AD011850	AD011850
C 603	20	1.6	786	5	BW079936	BW079936	676	20	1.5	346	2	AW363358	AW363358
C 604	20	1.6	789	8	BZ396217	EINAY14TF	677	20	1.5				
C 605	20	1.6	789	9	CG397286	ZMMBBc001	678	20	1.5				
C 606	20	1.6	792	5	BW142580	BW142580	C 679	20	1.5				
C 607	20	1.6	795	8	BZ452886	BONSXB0TR	C 680	20	1.5				
608	20	1.6	803	5	BU251348	603403612	C 681	20	1.5				

682	19	1.5	349	4	BJ378821	BJ378821	755	19	1.5	521	7	CN741425	SAL_US0004
C 683	19	1.5	360	7	D37065	D37065	756	19	1.5	521	8	AZ728536	RPCI-24--
C 684	19	1.5	360	9	C2250570	C2250570	757	19	1.5	525	4	BG565051	BG565051
C 685	19	1.5	362	4	BJ318349	BJ318349	C 758	19	1.5	526	6	CB720030	AMGNNUC:N
C 686	19	1.5	362	8	BH582883	BH582883	C 759	19	1.5	527	1	AU261975	AU261975
C 687	19	1.5	365	4	BM161546	BM161546	C 760	19	1.5	528	8	AQ776195	HS_3119_A
C 688	19	1.5	370	6	CD4117046	CD4117046	C 761	19	1.5	528	5	AZ074619	RPCI-23--4
C 689	19	1.5	371	8	A2953328	A2953328	C 762	19	1.5	529	8	EX563688	EX563688
C 690	19	1.5	374	2	BF549676	BF549676	C 763	19	1.5	529	5	CK295878	EST758592
C 691	19	1.5	376	1	A1551625	A1551625	C 764	19	1.5	536	7	CL587099	OB_BA008
C 692	19	1.5	387	5	BY024850	BY024850	C 765	19	1.5	536	9	AG201375	Pan trogl
C 693	19	1.5	388	1	A1781852	A1781852	C 766	19	1.5	538	9	CN113897	tigr-g88-
C 694	19	1.5	388	7	CV128874	CV128874	C 767	19	1.5	540	7	CN116152	ECSCAA44A
C 695	19	1.5	388	7	CV129655	CV129655	C 768	19	1.5	548	9	CR105821	Reverse s
C 696	19	1.5	390	4	BL181980	BL181980	C 769	19	1.5	548	9	CR105821	Reverse s
C 697	19	1.5	393	7	CK778830	CK778830	C 770	19	1.5	551	4	BJ341865	BJ341865
C 698	19	1.5	395	9	AG253558	AG253558	C 771	19	1.5	552	4	BI514156	BI514156
C 699	19	1.5	396	5	BU497167	BU497167	C 772	19	1.5	552	6	CB110335	CB110335
C 700	19	1.5	398	1	A1814774	A1814774	C 773	19	1.5	555	8	AQ490052	AQ490052
C 701	19	1.5	402	5	BU495641	BU495641	C 774	19	1.5	556	9	CR842873	CR842873
C 702	19	1.5	402	7	CK340612	CK340612	C 775	19	1.5	558	1	AJ500616	AJ500616
C 703	19	1.5	403	8	A2599537	A2599537	C 776	19	1.5	563	4	BI815498	BI815498
C 704	19	1.5	406	1	AA706466	AA706466	C 777	19	1.5	565	2	BF353126	BF353126
C 705	19	1.5	416	5	BP671122	BP671122	C 778	19	1.5	568	6	CB094370	CB094370
C 706	19	1.5	417	7	CV187916	CV187916	C 779	19	1.5	568	4	BI701369	BI701369
C 707	19	1.5	425	8	BH251978	BH251978	C 780	19	1.5	569	5	CE225595	CE225595
C 708	19	1.5	427	5	A1273387	A1273387	C 781	19	1.5	569	5	BU297731	BU297731
C 709	19	1.5	428	1	A4243632	A4243632	C 782	19	1.5	570	4	BI928878	BI928878
C 710	19	1.5	430	7	CF546295	CF546295	C 783	19	1.5	572	9	CE177344	CE177344
C 711	19	1.5	430	9	CE549393	CE549393	C 784	19	1.5	575	2	BE052275	BE052275
C 712	19	1.5	431	1	AV792057	AV792057	C 785	19	1.5	577	9	CK582894	CK582894
C 713	19	1.5	431	8	B91032	B91032	C 786	19	1.5	578	7	AW038487	AW038487
C 714	19	1.5	438										







Qy	1173	TGGCAATCAACAATTTTCTGGCCAGAGTGTTTTAAACTTCAGTTTCACAGAAAATAAGGCAC	1232
Db	961	TGGCAATCAACAATTTTCTGGCCAGAGTGTTTTAAACTTCAGTTTCACAGAAAATAAGGCAC	1020
Qy	1233	CCATCTGTCTGCCAACCTAAACACTCTTTTCGGTAGGTGGAAAGC	1274
Db	1021	CCATCTGTCTGCCAACCTAAACACTCTTTTCGGTAGGTGGAAAGC	1062

RESULT 2	CR607323	1360 bp	mrna	lineag	HTC 21-JUL-2004
LOCUS	CR607323	full-length cDNA clone CS0DC023YH14 of Neuroblastoma Cot			
DEFINITION	CR607323	25-normalized of Homo sapiens (human).			
ACCESSION	CR607323.1	GI:50488130			
VERSION	CR607323	HTC; CNSLT cDNA.			
KEYWORDS	CR607323.1	Homo sapiens (human)			
SOURCE	CR607323	Homo sapiens			
ORGANISM	CR607323	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	CR607323	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.			
AUTHORS	CR607323	1 (bases 1 to 1360)			
TITLE	CR607323	Full-length cDNA libraries and normalization			
JOURNAL	CR607323	Unpublished			
REMARK	CR607323	Contact : Feng Yiang Email : fliang@lifetech.com URL : <a href="http://fulllength.invitrogen.com/">http://fulllength.invitrogen.com/</a> Invitrogen Corporation 1600			

REFERENCE	TITLE	JOURNAL	COMMENT	FEATURES
2 (bases 1 to 1360)	Genoscope. Direct Submission			source
	Submitted (20-JUN-2004)	Genoscope - Centre National de Sequencage :		
	BP 191 91006 EVRY cedex - FRANCE	(E-mail : seqref@genoscope.cns.fr		
	- Web : www.genoscope.cns.fr)			
	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime			
	end enriched, double-strand cDNA was digested with Not I and cloned			
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library			
	was normalized. Library was constructed by Life Technologies, a			
	division of Invitrogen.			
	Location/Qualifiers			
	1. .1360			

617	QY	AAATATCGGCTTCTTGATGTGAAATAA	CTGCGCA	CACACAGAAAAAGCAGC	CAATGAAG	676
301	Db	AAATATCGCTTCTCTGATGTGAAATAA	CTGCGCA	CACACAGAAAAAGCAGC	CAATGAAG	360
677	QY	AGCTCTGTGAGTTT	AGGCGCTTATTCCTGT	CCGTGCCCTGGTGCTCT	CTGTTAAATGCAAG	736
361	Db	AGCTCTGTGAGTTT	AGGCGCTTATTCCTGT	CCGTGCCCTGGTGCTCT	CTGTTAAATGCGAAG	420
737	QY	GCTCTCTGATGCTGTAATGCCCCAT	CTGATGCATCAGCAT	TAAGTCCATTTACAACCTTAC	796	
421	Db	GCTCTCTGATGCTGTAATGCCCCAT	CTGATGCATCAGCAT	TAAGTCCATTTACAACCTTAC	480	
797	QY	AGGGAGAGGATTAAGTTTTTCTTGCT	TACAGACATTAATCTTCTCGTGTGCTGTGACTGG	856		
481	Db	AGGGAGAGGATTAAGTTTTTCTTGCT	TACAGACATTAATCTTCTCGTGTGCTGTGACTGG	540		
857	QY	TGATGATGCAGTCTGTGTTTTGGCTTT	CACTTCA	TGTTAGTCTTTAGAGAAACAGAAAAAT	916	
541	Db	TGATGATGCAGTCTGTGTTTTGGCTTT	CACTTCA	TGTTAGTCTTTAGAGAAACAGAAAAAT	600	
917	QY	ACGATGGTCACGACGATGTTCTTCGCAAT	CGTACAGCTGATAGGAA	CACGGAAGCAAGCTG	976	
601	Db	ACGATGGTCACGACGATGTTCTTCGCAAT	CGTACAGCTGATAGGAA	CACGGAAGCAAGCTG	660	
977	QY	AAAAATTTGCTTACCAGCTTGAGCTAAAT	GGTGCATAGGCGACGATGACTTGGGGAAGCGA	1036		
661	Db	AAAAATTTGCTTACCAGCTTGAGCTAAAT	GGTGCATAGGCGACGATGACTTGGGGAAGCGA	720		
1037	QY	CTCTCGATCTATTCTAGAGAAAT	TGCAACGCCATTAATGAATAGCGACTGTCTAGTCT	1096		
721	Db	CTCTCGATCTATTCTAGAGAAAT	TGCAACGCCATTAATGAATAGCGACTGTCTAGTCT	780		
1097	QY	TTGACACACGCAATTCACAGCTTTTTT	TGAGAAATGGCAATTTAGGCATCAATGTAAC	1156		
781	Db	TTGACACACGCAATTCACAGCTTTTTT	TGAGAAATGGCAATTTAGGCATCAATGTAAC	840		
1157	QY	TTTCCATGTGTTGAAATGCAATCAAA	CAATTTCTGCGCAGTGTGTTAAAACTTCAGTTTC	1216		
841	Db	TTTCCATGTGTTGAAATGCAATCAAA	CAATTTCTGCGCAGTGTGTTAAAACTTCAGTTTC	900		
1217	QY	ACAGAAAAATAGGCCACCCATCTGTG	TGCGCAACTAAAACTCTTTTCGGTAGTGGGAAGC	1274		
901	Db	ACAGAAAAATAGGCCACCCATCTGTG	TGCGCAACTAAAACTCTTTTCGGTAGTGGGAAGC	958		

RESULT 3	CR614925	1857 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	CR614925				
DEFINITION	full-length cDNA clone CS0DK012Y120 of HeLa cells Cot 25-normalized of Homo sapiens (human) .				
ACCESSION	CR614925				
VERSION	CR614925.1	GI:50495732			
KEYWORDS	HTC; CNSLT_cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1857)				
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
REMARK	Unpublished				
	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InvitroGen Corporation 1600 Faraday Avenue				
REFERENCE	2 (bases 1 to 1857)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library				



```
Db 301 CTGTGAGTTTGGGCTTATCTCTGTGCGTCCCTGGTGTCTCTCTGTAAATGGCAGGCTC 360
Qy 741 TCTGGATGCTGTATATGCCCCATCTGATGATCATAGCATAAGTCCATTACAAACCTCAGAGG 800
Db 361 TCTGGATGCTGTATATGCCCCATCTGATGATCATAGCATAAGTCCATTACAAACCTCAGAGG 420
Qy 801 AGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGAT 860
Db 421 AGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGAT 480
Qy 861 GATGCACTCTGTTTGGCTTTCACATTCATGTTAGTCTTAGAGAAACAGGAAATATACGA 920
Db 481 GATGCACTCTGTTTGGCTTTCACATTCATGTTAGTCTTAGAGAAACAGGAAATATACGA 540
Qy 921 TGCTCACAGAGCTTCTTCCCAATCTGATAGTCTGATAGGAAACAGGAAATATACGA 980
Db 541 TGCTCACAGAGCTTCTTCCCAATCTGATAGTCTGATAGGAAACAGGAAATATACGA 600
Qy 981 TTTTGGCTTACCGACTTGGCTAAATGGTCTATAGCGAGCAATTCAGTTGGGAAGCGACTCC 1040
Db 601 TTTTGGCTTACCGACTTGGCTAAATGGTCTATAGCGAGCAATTCAGTTGGGAAGCGACTCC 660
Qy 1041 TCGATCTATTTCATGAAGGAATTCACAGCAATTCATGATAGGAGCTGTCTAGTCTTTTGA 1100
Db 661 TCGATCTATTTCATGAAGGAATTCACAGCAATTCATGATAGGAGCTGTCTAGTCTTTTGA 720
Qy 1101 CACCAGCATTGCAACAGCTTTTTCAGAGAAATGCAATTTAGGATCAATGTAATTTTC 1160
Db 721 CACCAGCATTGCAACAGCTTTTTCAGAGAAATGCAATTTAGGATCAATGTAATTTTC 780
Qy 1161 CATGCTTGAATGGCAATCAACATTTTCGCCAGTGTGTTAACTTCAGTTTCACAG 1220
Db 781 CATGCTTGAATGGCAATCAACATTTTCGCCAGTGTGTTAACTTCAGTTTCACAG 840
Qy 1221 AAAATAAGCACCACATCTGTCTGCCAACCTTAACTCTTTCGTTAGTGTGAGC 1274
Db 841 AAAATAAGCACCACATCTGTCTGCCAACCTTAACTCTTTCGTTAGTGTGAGC 894
```

```
RESULT 5
AY402255
LOCUS
DEFINITION Homo sapiens SIAH1 gene, VIRUAL TRANSCRIPT, partial sequence, 849 bp DNA linear GSS 15-DEC-2003
ACCESSION AY402255
VERSION AY402255.1 GI:39758241
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 849)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1950-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 849)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
location/Qualifiers
1..849
source
```

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>849
/gene="SIAH1"
/locus_tag="HCM1163"
```

## ORIGIN

```
Query Match 66.6%; Score 849; DB 9; Length 849;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 849; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 322 ATGAGCGCTCAGACTCTACAGCAATTTACCTACCGGTACCTCGAAGTGTCCACCATCCAG 381
Db 1 ATGAGCGCTCAGACTCTACAGCAATTTACCTACCGGTACCTCGAAGTGTCCACCATCCAG 60
Qy 382 AGGGTGCCTGCCCTGACTGGCACTGCAATCAACAAATGACTTGGCGAGTCTTTTGGAG 441
Db 61 AGGGTGCCTGCCCTGACTGGCACTGCAATCAACAAATGACTTGGCGAGTCTTTTGGAG 120
Qy 442 TGTCCAGTCTGCTTTGACTATGTGTTACCGCCCAATTTCAATGTGAGAGTGGCCATCTT 501
Db 121 TGTCCAGTCTGCTTTGACTATGTGTTACCGCCCAATTTCAATGTGAGAGTGGCCATCTT 180
Qy 502 GTTTGTAGCACTGTGCGCCCAAGCTCAGATGTTTCCAACTTGCCTGGGGCCCTTTGGGA 561
Db 181 GTTTGTAGCACTGTGCGCCCAAGCTCAGATGTTTCCAACTTGCCTGGGGCCCTTTGGGA 240
Qy 562 TCCATTTCGCACTTGGCTATGAGAAAGTGGCTTAATTCAGTACTTTTCCCTGTAATAT 621
Db 241 TCCATTTCGCACTTGGCTATGAGAAAGTGGCTTAATTCAGTACTTTTCCCTGTAATAT 300
Qy 622 GCGTCTTCTGGATGTGAATAAATCTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTC 681
Db 301 GCGTCTTCTGGATGTGAATAAATCTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTC 360
Qy 682 TGTGAGTTTAGGCTTATTTCTGTCGCTGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCT 741
Db 361 TGTGAGTTTAGGCTTATTTCTGTCGCTGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCT 420
Qy 742 CTGGATGCTGTAATGCCCATCTGATGCAATGAGCAATGAGTCCATTCACACCTTACAGGGA 801
Db 421 CTGGATGCTGTAATGCCCATCTGATGCAATGAGCAATGAGTCCATTCACACCTTACAGGGA 480
Qy 802 GAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTG 861
Db 481 GAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTG 540
Qy 862 ATGCACTCTGTTTGGCTTTCACATTCATGTTAGTCTTACAGAAACAGGAAAAATACGAT 921
Db 541 ATGCACTCTGTTTGGCTTTCACATTCATGTTAGTCTTACAGAAACAGGAAAAATACGAT 600
Qy 922 GGTACACAGCAGTCTTCTGCAATCTGATGATAGGAAACAGGAAACAGGAAACAGGAAAT 981
Db 601 GGTACACAGCAGTCTTCTGCAATCTGATGATAGGAAACAGGAAACAGGAAACAGGAAAT 660
Qy 982 TTTGCTTACCGACTTGGCTAAATGCTCATAGCGCAATTCAGTCTGGAAGCGACTCTCT 1041
Db 661 TTTGCTTACCGACTTGGCTAAATGCTCATAGCGCAATTCAGTCTGGAAGCGACTCTCT 720
Qy 1042 CGATCTATTCATGAAGGAATTGCAACAGCCATTTATGAATAGCGACTGTCTAGTCTTTGAC 1101
Db 721 CGATCTATTCATGAAGGAATTGCAACAGCCATTTATGAATAGCGACTGTCTAGTCTTTGAC 780
Qy 1102 ACCAGCATGACAGCTTTTTCAGAAAAATGCAATTTAGGCAATCAATGTAATATTTCC 1161
Db 781 ACCAGCATGACAGCTTTTTCAGAAAAATGCAATTTAGGCAATCAATGTAATATTTCC 840
Qy 1162 ATGTGTTGA 1170
Db 841 ATGTGTTGA 849
```

```
RESULT 6
CR593197      1576 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS      full-length cDNA clone CS0D1011Y112 of Placenta Cot 25-normalized
DEFINITION      of Homo sapiens (human).
ACCESSION      CR593197
VERSION      CR593197.1 GI:50474004
KEYWORDS      HTC; CNSLT_CDNA.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      1 (bases 1 to 1576)
JOURNAL      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
REMARK      Full-length cDNA libraries and normalization
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1576)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope, Centre National de Sequencage :
BP 191 91006 EVRY Cedex - FRANCE [E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr]
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1011Y112"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT 6"
ORIGIN
Query Match 53.8%; Score 685; DB 3; Length 1576;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 685; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 590 TGGCTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCGATGTGAATTAACCTCTGC 649
DB 1 TGGCTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCGATGTGAATTAACCTCTGC 60
QY 650 CACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCTTATTCCTGTCCT 709
DB 61 CACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCTTATTCCTGTCCT 120
QY 710 GCCTGGTGCTTCTGTAATGCGAAGCTCTCTGATGCTGTATGCGCCATCTGATGC 769
DB 121 GCCTGGTGCTTCTGTAATGCGAAGCTCTCTGATGCTGTATGCGCCATCTGATGC 180
QY 770 ATCAGCATAAGTCCATTACACCCCTACAGGAGGATATAGTTTCTGTCTACAGACA 829
DB 181 ATCAGCATAAGTCCATTACACCCCTACAGGAGGATATAGTTTCTGTCTACAGACA 240
QY 830 TTAATCTTCTGGTGTCTGTGATCGGGTGATGATGATGATGATGATGATGATGATGAT 889
DB 241 TTAATCTTCTGGTGTCTGTGATCGGGTGATGATGATGATGATGATGATGATGATGAT 300
QY 890 TGTAGTCTTAGAAGAACAGGAAGAAATACATGGTCCACAGAGTCTTCGCAATCGTAC 949
DB 301 TGTAGTCTTAGAAGAACAGGAAGAAATACATGGTCCACAGAGTCTTCGCAATCGTAC 360
QY 950 AGCTGATAGGAACACCAAGCAGCTGAAAATTTTGTCTTACCGACTTGAGCTAAATGGTC 1009
DB 361 AGCTGATAGGAACACCAAGCAGCTGAAAATTTTGTCTTACCGACTTGAGCTAAATGGTC 420
QY 1010 ATAGGCGCAGATTGACTTGGGAAGCGACTCCTCGATCTATTTCATGAAGGAATGCAACAG 1069
```

```
Db 421 ATAGCGCAGGATGACTTGGGAAGCGACTCCTCGATCTATTTCATGAGGAATTCGAACAG 480
QY 1070 CCATTTCGAATAGGAGCTGTCTAGTCTTTGACACCGACCATTCGACAGCTTTTTCGAGAAA 1129
DB 481 CCATTTCGAATAGGAGCTGTCTAGTCTTTGACACCGACCATTCGACAGCTTTTTCGAGAAA 540
QY 1130 ATGGCAATTTAGGCAATCAATGTAATCTATTTCCATGTTTGAATGGCAATCAAAACATTTT 1189
DB 541 ATGGCAATTTAGGCAATCAATGTAATCTATTTCCATGTTTGAATGGCAATCAAAACATTTT 600
QY 1190 CTGGCCAGCTGTTTAAACTTCAGTTTCACAGAAATAAGGACCCCATCTGTCGCAAC 1249
DB 601 CTGGCCAGCTGTTTAAACTTCAGTTTCACAGAAATAAGGACCCCATCTGTCGCAAC 660
QY 1250 TAAACTCTTTTGGGTAGGTGAAGC 1274
DB 661 TAAACTCTTTTGGGTAGGTGAAGC 685

RESULT 7
BQ447177      714 bp      mRNA      linear      EST 29-MAY-2002
LOCUS      UI-H-EUI-Bad-i-20-0-UI.s1 NCI_CGAP_Ctl Homo sapiens cDNA clone
DEFINITION      UI-H-EUI-Bad-i-20-0-UI 3', mRNA sequence.
ACCESSION      BQ447177
VERSION      BQ447177.1 GI:21250289
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      1 (bases 1 to 714)
JOURNAL      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Straubeberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-53, >POLY A#Simple_repeat (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers
1..714
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EUI-Bad-i-20-0-UI"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Ctl"
/note="Organ: Knee; Vector: pUT3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI CGAP Ctl is a normalized cDNA library containing the
following tissue(s): Osteoarthritic Cartilage The library
was constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pUT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
TGATCAGCT.
```





Accession	Position	Sequence	Length
Db	326	GCAAGCTGAAAATTTGCTTACCAGATTGAGCTTAAATGGTCTAGGCGACGATTGACTTG	267
Qy	1029	GGAAGCGACTCCTCGATCTATTTCATCAAGGAATTGCAACAGCCATTATGAATAGCGACTG	1088
Db	266	GGAAGCGACTCCTCGATCTATTTCATCAAGGAATTGCAACAGCCATTATGAATAGCGACTG	207
Qy	1089	TCTAGTCTTTTGACACCGACGATTGCACAGCTTTTTCAGAAAATGGCAATTTTAGGCATCAA	1148
Db	206	TCTAGTCTTTTGACACCGACGATTGCACAGCTTTTTCAGAAAATGGCAATTTTAGGCATCAA	147
Qy	1149	TGTAACTATTTTCATCTGTGTGAATGGCAATCAACATTTTCTGGCCAGTGTTTAAACT	1208
Db	146	TGTAACTATTTTCATCTGTGTGAATGGCAATCAACATTTTCTGGCCAGTGTTTAAACT	87
Qy	1209	TCAGTCTTCAGAGAAATTAAGGCCACCATCTGTCTGCCAACCTAAAACTCTTTCGGTAGGT	1268
Db	86	TCAGTCTTCAGAGAAATTAAGGCCACCATCTGTCTGCCAACCTAAAACTCTTTCGGTAGGT	27
Qy	1269	GGAAGC	1274
Db	26	GGAAGC	21

RESULT 9	849 bp	DNA	linear	GSS 15-DEC-2003
AY402256				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
PUBMED				
COMMENT				
FEATURES				
source				
gene				
ORIGIN				
Query Match	50.6%	Score 645;	DB 9;	Length 849;
Best Local Similarity	99.5%	Pred. No. 0;		
Matches	845;	Conservative 0;	Mismatches 4;	Indels 0;
Gaps	0;			
322	ATGAGCCGTACAGCTGCTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCGAG	381		
1	ATGAGCCGTACAGCTGCTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCGAG	60		



TITLE  
JOURNAL  
COMMENT

HRI-human cDNA project  
Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1530-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

## FEATURES

source

Location/Qualifiers  
1. 812  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HEMBA1006927"  
/tissue\_type="whole embryo, mainly head"  
/dev\_stage="embryo, 10 weeks"  
/clone\_lib="HEMBA1"  
/note="Vector: pME18SFL3"

## ORIGIN

Query Match 50.4%; Score 642; DB 1; Length 812;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 633 ATGTGAAATAAATCTGCCACACACAGAAAGGACACCATGAGAGCTCTGTGAGTTTAA 692  
DB 800 ATGTGAAATAAATCTGCCACACACAGAAAGGACACCATGAGAGCTCTGTGAGTTTAA 741  
QY 693 GCCTTATTCCTGTCGCGCCCTGCTCTCTGTAATGCGAAGGCTCTCTGATGCTCT 752  
DB 740 GCCTTATTCCTGTCGCGCCCTGCTCTCTGTAATGCGAAGGCTCTCTGATGCTCT 681  
QY 753 AATGCCCATCTGATGATCAGCATAGTCCATTACCAAGCTCTACAGGGAGAGATATAGT 812  
DB 680 AATGCCCATCTGATGATCAGCATAGTCCATTACCAAGCTCTACAGGGAGAGATATAGT 621  
QY 813 TTTTCTGTACAGACATTAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 872  
DB 620 TTTTCTGTACAGACATTAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561  
QY 873 TTTTGTGCTTTCACTTCATGTTAGTCTTTAGAGAAACAGGAAATACGATGGTCACAGCA 932  
DB 560 TTTTGTGCTTTCACTTCATGTTAGTCTTTAGAGAAACAGGAAATACGATGGTCACAGCA 501  
QY 933 GTTCTTGGCAATCGTACAGCTGATAGGAAACAGCAAGCAAGCTGAAATTTTGTTCACG 992  
DB 500 GTTCTTGGCAATCGTACAGCTGATAGGAAACAGCAAGCAAGCTGAAATTTTGTTCACG 441  
QY 993 ACTTGAGCTTAATGTCATAGGAGCAGTATGCTGGGAGGAGCTCTGCTGATCTATTCA 1052  
DB 440 ACTTGAGCTTAATGTCATAGGAGCAGTATGCTGGGAGGAGCTCTGCTGATCTATTCA 381  
QY 1053 TGAAGGAATTCACACAGCATTATTAATAGCAGCTGCTGAGTCTTTGACACCGAGATTGC 1112  
DB 380 TGAAGGAATTCACACAGCATTATTAATAGCAGCTGCTGAGTCTTTGACACCGAGATTGC 321  
QY 1113 ACAGCTTTTTCAGAAAAATGGCAATTTAGGCATCAATGTAATCTATTTCATGTTGAAA 1172  
DB 320 ACAGCTTTTTCAGAAAAATGGCAATTTAGGCATCAATGTAATCTATTTCATGTTGAAA 261  
QY 1173 TGGCAATCAAACTTTTCGGCAGTGTATAAACTTCAGTTTCACAGAAAAATAGGCAC 1232  
DB 260 TGGCAATCAAACTTTTCGGCAGTGTATAAACTTCAGTTTCACAGAAAAATAGGCAC 201  
QY 1233 CCATCTGTCTGCAACCTTAACCTCTTTTCGGTAGTGGAGC 1274  
DB 200 CCATCTGTCTGCAACCTTAACCTCTTTTCGGTAGTGGAGC 159

## RESULT 11

BX366404  
LOCUS BX366404 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
DEFINITION CDNA clone CSODC023YH14 5-PRIME, mRNA sequence.  
ACCESSION BX366404  
VERSION BX366404.2 GI:46572835  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1. (bases 1 to 859)  
AUTHORS Li, W.B., Grubel, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2000)  
COMMENT On May 8, 2003 this sequence version replaced gi:30447532.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
8104.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdnap?S=CSODC023DD07QP2&amp;c=8104.r.

## FEATURES

source

Location/Qualifiers

1. 859  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODC023YH14"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 50.0%; Score 637; DB 5; Length 859;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 687; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 317 CAGAAATGAGCGCTCAGACTGCTACAGCAATTACCTACCGGTACCTCGAAGTGTCCACCAT 376  
DB 1 CAGAAATGAGCGCTCAGACTGCTACAGCAATTACCTACCGGTACCTCGAAGTGTCCACCAT 60  
QY 377 CCCAGAGGTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436  
DB 61 CCCAGAGGTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
QY 437 TTGAGTGTCCAGTCTGCTTTGACTATGTTGTTACCGCCCATCTTCAATGTCAGAGTGGCC 496  
DB 121 TTGAGTGTCCAGTCTGCTTTGACTATGTTGTTACCGCCCATCTTCAATGTCAGAGTGGCC 180  
QY 497 ATCTTTGTTGTAGCAACTGTGCGCCAAAGCTCACAATGTTGTCCAACTTGCCGGGGCCCTT 556  
DB 181 ATCTTTGTTGTAGCAACTGTGCGCCAAAGCTCACAATGTTGTCCAACTTGCCGGGGCCCTT 240  
QY 557 TGGGATTCATTCGCAACTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTGA 616  
DB 241 TGGGATTCATTCGCAACTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTGA 300  
QY 617 AATATGCGCTTCTCGATGTGAATAAATCTTCCACACACAGAAAAAGCAGACCATGAG 676  
DB 301 AATATGCGCTTCTCGATGTGAATAAATCTTCCACACACAGAAAAAGCAGACCATGAG 360  
QY 677 AGCTCTGTGAGTTTAGGCGCTTATCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736

361	Db		AGCTCTGTGAGTTTAGGCTTAATTCCTGTCGGTGCCTGTGTCTTCTGTAAATGCGAAG	420
737	Qy		GCTCTCTGATGCTGTAATGCCCCATCTGATGCAATCAGCATATAAGTCCATTACAAACCCCTAC	796
421	Db		GCTCTCTGGAGCTGTTATGCCCCATCTGATGCATCAGCATAGTCCATTACAAACCCCTAC	480
797	Qy		AGGGAGAGGATATAGTGTTTTCTTGCTACAGACATTAATCTTCTGTCGTGTGACTGGG	856
481	Db		AGGGAGAGGATATAGTGTTTTCTTGCTACAGACATTAATCTTCTGTCGTGTGACTGGG	540
857	Qy		TGATGATGCAGTCTCTGTTTTGGCTTTTCACTTCATGTTTAGTCTTTAGAGAACAGGAAAAAT	916
541	Db		TGATGATGCAGTCTCTGTTTTGGCTTTTCACTTCATGTTTAGTCTTTAGAGAACAGGAAAAAT	600
917	Qy		ACGATGGTCCACGACGAGTCTTTTCGCAATCGTACAGCTGATAGGAACAACGACAGCAAGCTG	976
601	Db		ACGATGGTCCACGACGAGTCTTCGCAATCGTACAGCTGATAGGAACAACGACAGCAAGCTG	660
977	Qy		AAAAATTTTGGCTTACCGACTTGAGCTAAA	1004
661	Db		AAAAATTTTGGCTTACCGACTTGAGCTAAA	688

RESULT 12	
BU6622334	
LOCUS	
DEFINITION	
ACCESSION	BU6622334
VERSION	BU6622334.1
KEYWORDS	GI:23374513
SOURCE	Est.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (bases 1 to 693)
COMMENT	Gubin,A.N., Lee,Y.T., Bouffard,G.G. and Miller,J.L. Gene Expression in Human Erythroid Precursor Cells Unpublished (2002) Contact: Jeffery L. Miller Laboratory of Chemical Biology National Institute of Diabetes and Digestive and Kidney Diseases Building 10, Room 9B17, National Institutes of Health, Bethesda, MD 20892, USA Tel: 301 402 2373

```

FEATURES
source
seq primer: 5'-taccata-cttctgacx sequencing primer.
Location/Qualifiers
1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cl84e12"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Precursor Cells"
/cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
/dev_stage="Precursor erythroblasts; GPA++"
/lab_host="DH5alpha"
/clone_lib="Hembase; Erythroid Precursor Cells (ICB:cl
library)"
/note="Organ: blood; Vector: pTriplex2; Site 1: SfiI;
Site 2: SfiI; A complementary DNA (cDNA) library from
human erythroid precursor cells was constructed using

```

SMART PCR (polymerase chain reaction) cDNA Library Construction Kit (Clontech, Palo Alto, CA) according to the manufacturer's directions, but with slight modifications. Briefly, reverse transcription was performed in the presence of 1  $\mu$ mol/L peptide nucleic acid (PNA) oligos (N'-terminal)-biotin-GTC-CAC-CCG-ARG-CYT-G- (C-terminal) and (N'-terminal)-biotin-C(T/C)-T-GAA-GTT-CTC-AGG-A- (C-terminal) . Synthesized cDNA was digested with SfiI and size-selected on a 1% agarose gel (>800bp). Large-scale sequencing of the library was performed by the NIH Intramural Sequencing Center (NISC; <http://www.nisc.nih.gov/>).

ORIGIN	Query Match	49.7%;	Score 633;	DB 5;	Length 693;				
	Best Local Similarity	99.9%;	Pred. No. 0;						
	Matches 683;	Conservative	0;	Mismatches 1;	Indels 0;				
	Gaps	0;							
QY	95	GGAA	CA	TTT	TGAAAAGAGAGCTATCCAGGTGACAGATCCTAATAAAGTGCAATTCAGT	154			
DB	10	GGAA	CA	TTT	TGAAAAGAGAGCTATCCAGGTGACAGATCCTAATAAAGTGCAATTCAGT	69			
QY	155	GTA	ATTT	TA	TTTTT	TAATATCTTTTTTAATCCTATTTTTTCTTCCTCTTTTGTCTCAGTAA	214		
DB	70	GTA	ATTT	TA	TTTTT	TAATATCTTTTTTAATCCTATTTTTTCTTCCTCTTTTGTCTCAGTAA	129		
QY	215	ATTT	TG	TAT	GAAAC	TTTTT	AAAAAGACCTTATGGGATGTAAACATTTATATAAAGTAAGTCA	274	
DB	130	ATTT	TG	TAT	GAAAC	TTTTT	AAAAAGACCTTATGGGATGTAAACATTTATATAAAGTAAGTCA	189	
QY	275	TGG	TTA	TAA	TAT	TTTTT	TCCCTGCCCTCCTATGTATTTATTTTTCAGAAATGAGCCGTGCA	334	
DB	190	TGG	TTA	TAA	TAT	TTTTT	TCCCTGCCCTCCTATGTATTTATTTTTCAGAAATGAGCCGTGCA	249	
QY	335	CTG	CT	AC	AG	CA	TTT	ACCTCGAGTGTCCACATCCACAGAGGGTGCCTGCCCC	394
DB	250	CTG	CT	AC	AG	CA	TTT	ACCTCGAGTGTCCACATCCACAGAGGGTGCCTGCCCC	309
QY	395	TG	ACT	GG	CA	CA	CT	GTGCGAGTCTTTTGTAGTGTCCAGTCTGCT	454
DB	310	TG	ACT	GG	CA	CA	CT	GTGCGAGTCTTTTGTAGTGTCCAGTCTGCT	369
QY	455	TTG	AC	TAT	GT	TAC	CGCCCAT	CTTCAATGTCAGAGTGCCCATCTTGTTGTGTAGCAACT	514
DB	370	TTG	AC	TAT	GT	TAC	CGCCCAT	CTTCAATGTCAGAGTGCCCATCTTGTTGTGTAGCAACT	429
QY	515	GT	CG	CC	CA	AA	GT	GTCCAACTTTGCGGGGGCCCTTTGGGATCCATTCGCAACT	574
DB	430	GT	CG	CC	CA	AA	GT	GTCCAACTTTGCGGGGGCCCTTTGGGATCCATTCGCAACT	489
QY	575	TG	CT	AT	TG	AG	AA	GTGGCTAATCTTTTGCCTGTATAATATGCGTCTTCTGGAT	634
DB	490	TG	CT	AT	TG	AG	AA	GTGGCTAATCTTTTGCCTGTATAATATGCGTCTTCTGGAT	549
QY	635	GT	GA	AT	TA	CT	TG	CAACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGC	694
DB	550	GT	GA	AT	TA	CT	TG	CAACACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGC	609
QY	695	CTT	ATT	CT	CT	GT	CG	CCCTGGTCTTCTGTAAATGGCAAGGCTCTCTGGATGCTGTAA	754
DB	610	CTT	ATT	CT	CT	GT	CG	CCCTGGTCTTCTGTAAATGGCAAGGCTCTCTGGATGCTGTAA	669
QY	755	TG	CC	CA	AT	CT	GT	ATCAGCAT	778
DB	670	TG	CC	CA	AT	CT	GT	ATCAGCAT	693

RESULT 13	LOCUS	DEFINITION
BG619351		

BG619351 824 bp mRNA linear EST 18-APR-2001  
 602619484F1 NIH\_MGC\_79 Homo sapiens cDNA clone IMAGE:4733228 5',  
 mRNA sequence.

ACCESSION	BC619351	GI:13670722
VERSION	BC619351.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	1 (bases 1 to 824)	
REFERENCE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .	
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)	
TITLE	Unpublished (1999)	
JOURNAL	Contact: Robert Strausberg, Ph.D.	
COMMENT	Email: <a href="mailto:cgabbs@mail.nih.gov">cgabbs@mail.nih.gov</a>	
	Tissue Procurement: CLONTECH Laboratories, Inc.	
	CDNA Library Preparation: CLONTECH Laboratories, Inc.	
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
	DNA Sequencing by: Incyte Genomics, Inc.	
	Clone distribution: MGC clone distribution information can be	
	found through the I.M.A.G.E. Consortium/LLNL at:	
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>	
	Plate: LLCM1594 row: g column: 21	
	High quality sequence spot: 748.	

## FEATURES

## ORIGIN

Query Match	49.0%;	Score 624;	DB 4;	Length 824;	
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches	674;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	400	GGCACAACTGCATCCCAACAAATGACTGGCGAGTCTTTTGGATGTCACAGTCTGCTTTGCAC	459		
DB	1	GGCACAACTGCATCCCAACAAATGACTGGCGAGTCTTTTGGATGTCACAGTCTGCTTTGCAC	60		
QY	460	TATGTGTTACCGCCCATCTTTCAATGTCAGAGTGGCCATCTTGTTGTGAGCAACTGTGCGC	519		
DB	61	TATGTGTTACCGCCCATCTTTCAATGTCAGAGTGGCCATCTTGTTGTGAGCAACTGTGCGC	120		
QY	520	CCAAAGCTCACATGTTGTCCAACTTGCGGGGGCCCTTTGGGATCCATTCGCAACTTGCT	579		
DB	121	CCAAAGCTCACATGTTGTCCAACTTGCGGGGGCCCTTTGGGATCCATTCGCAACTTGCT	180		
QY	580	ATGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATATGCGTCTCTCTGGATGTGAA	639		
DB	181	ATGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATATGCGTCTCTCTGGATGTGAA	240		
QY	640	ATAACTCTGCCACACACAGAAAAAGCAGACCATCAAGAGCTCTGTGAGTTTAGGCCCTTAT	699		
DB	241	ATAACTCTGCCACACACAGAAAAAGCAGACCATCAAGAGCTCTGTGAGTTTAGGCCCTTAT	300		
QY	700	TCCTGTCCGTGCCCTGGTGTCTTCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCC	759		
DB	301	TCCTGTCCGTGCCCTGGTGTCTTCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCC	360		
QY	760	CATCTGTAGTCATCAGCAATAAGTCATTACAACCCCTACAGGAGAGGATATAGTTTTCCTT	819		





insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH MGC Library."

ORIGIN	Query Match Best Local Similarity Matches 590; Conservative 0;	46.3%; Score 590; DB 4; Length 735; 100.0%; Pred. No. 2.3e-305; Matches 590; Mismatches 0; Indels 0; Gaps 0;
QY	317	CAGAAATGAGCGCTGACACTGCTACACGATTTACCTACCGGTACCTCGAAGTGTCCACCAT 376
DB	146	CAGAAATGAGCGCTGACACTGCTACACGATTTACCTACCGGTACCTCGAAGTGTCCACCAT 205
QY	377	CCCAGAGGGTGCCTGCCCTGACTGGCAACAATGTCATCCAAACAATGACTTGGCGAGTCTTT 436
DB	206	CCCAGAGGGTGCCTGCCCTGACTGGCAACAATGTCATCCAAACAATGACTTGGCGAGTCTTT 265
QY	437	TTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCATCTTCAATGTCAGATGGCC 496
DB	266	TTGAGTGTCCAGTCTGCTTTGACTATGTGTTACCGCCCATCTTCAATGTCAGATGGCC 325
QY	497	ATCTTGTGTTGTAGCAACTGTGCGCCAAAGCTCACATGTGTCCAACTTGCGGGGGCCCTT 556
DB	326	ATCTTGTGTTGTAGCAACTGTGCGCCAAAGCTCACATGTGTCCAACTTGCGGGGGCCCTT 385
QY	557	TGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTTAATTGAGTACTTTTCCCTGTGA 616
DB	386	TGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTTAATTGAGTACTTTTCCCTGTGA 445
QY	617	AATATGCGTCTCTGGAGTGTGAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAG 676
DB	446	AATATGCGTCTCTGGAGTGTGAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAG 505
QY	677	AGCTCTGTGAGTTTAGCGCTTATTCCTGTCGGTCCCTGGTGCTTCCGTAAATGGCAAG 736
DB	506	AGCTCTGTGAGTTTAGCGCTTATTCCTGTCGGTCCCTGGTGCTTCCGTAAATGGCAAG 565
QY	737	GCTCTCTGGATGCTGTAATGCCCCATCTGATGCAATCAGCATAGTCCATACACCCCTAC 796
DB	566	GCTCTCTGGATGCTGTAATGCCCCATCTGATGCAATCAGCATAGTCCATACACCCCTAC 625
QY	797	AGGAGAGGATATAGTTTTTCTTGCTACAGACATAATCTTCCTGGTGTGTTGACTGGG 856
DB	626	AGGAGAGGATATAGTTTTTCTTGCTACAGACATAATCTTCCTGGTGTGTTGACTGGG 685
QY	857	TGATGATGCACTCTGTTTTGGCTTTTCACTTCATGTTAGTCTTAGAGAA 906
DB	686	TGATGATGCACTCTGTTTTGGCTTTTCACTTCATGTTAGTCTTAGAGAA 735

RESULT 18	AI9711258	778 bp	mrna	linear	EST 08-MAR-2000
LOCUS	AI9711258				
DEFINITION	w27b08.x1 NCICGAP Pr28 Homo sapiens cDNA clone IMAGE:248887 3', similar to TR:Q43269 O43269 HSI41.1; mRNA sequence.				
ACCESSION	AI9711258				
VERSION	AI9711258.1	GI:5768084			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 778)				
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.				







library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoT of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

ORIGIN	Query Match	39.4%;	Score 502;	DB 7;	Length 586;	
	Best Local Similarity	100.0%;	Pred. No. 5e-258;			
	Matches 502;	Conservative	0;	Mismatches	0;	Gaps
QY	317	CAGAAATGAGCCGTGCTGCTACAGCAAT	TACTACCGGTACTCTCGAAGTGCCACCAT	376		
DB	85	CAGAAATGAGCCGTGCTGCTACAGCAAT	TACTACCGGTACTCTCGAAGTGCCACCAT	144		
QY	377	CCGAGAGGGTGCCTGCCCTGACTGGCAAC	TGCATCCAAATGACTTGGCGAGTCTTT	436		
DB	145	CCGAGAGGGTGCCTGCCCTGACTGGCAAC	TGCATCCAAATGACTTGGCGAGTCTTT	204		
QY	437	TTGAGTGTCAAGTCTGCTTTTGA	CTATGTGTTCACGCCCATCTTCAATGTCAGATGGCC	496		
DB	205	TTGAGTGTCAAGTCTGCTTTTGA	CTATGTGTTCACGCCCATCTTCAATGTCAGATGGCC	264		
QY	497	ATCTTGTGTTTACCACTGTGCGCCCAAG	CTCACATGTTGCCAACTTGCAGTGGGGGCCCTT	556		
DB	265	ATCTTGTGTTTACCACTGTGCGCCCAAG	CTCACATGTTGCCAACTTGCAGTGGGGGCCCTT	324		
QY	557	TGGGATCCATTTCGCAACTTGGCTATG	GAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTGA	616		
DB	325	TGGGATCCATTTCGCAACTTGGCTATG	GAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTGA	384		
QY	617	AATAATGCGTCTTTGGATGTGAAATAA	CTCTGGCACACACAGAAAAAGCAGACCATGAAG	676		
DB	385	AATAATGCGTCTTTGGATGTGAAATAA	CTCTGGCACACACAGAAAAAGCAGACCATGAAG	444		
QY	677	AGCTCTGTGAGTTTAGCCCTTAATTCCT	GTCGTCGCTGGTCTTCCGTAAATGGCAAG	736		
DB	445	AGCTCTGTGAGTTTAGCCCTTAATTCCT	GTCGTCGCTGGTCTTCCGTAAATGGCAAG	504		
QY	737	GCTCTCTGGATGCTGTAATGCCCATCT	GATGCATCAGCAATAGTCCATTACAACCCCTAC	796		
DB	505	GCTCTCTGGATGCTGTAATGCCCATCT	GATGCATCAGCAATAGTCCATTACAACCCCTAC	564		
QY	797	AGGAGAGGATATAGTTTTTCT	818			
DB	565	AGGAGAGGATATAGTTTTTCT	586			

RESULT 21				
AUI38740	AU138740	734 bp	mRNA linear	EST 02-AUG-2002
LOCUS				
DEFINITION	AUI38740	PLAC1 Homo sapiens cDNA clone PLACE109197 5', mRNA sequence.		

ACCESSION	AU138740	
VERSION	AU138740.1	GI:11000261
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

ORGANISM      Homo sapiens      Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**REFERENCE**

1 (bases 1 to 734)

**AUTHORS**

Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isegai, T.

TITLE HRI human cDNA project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu,  
Chiba

Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human CDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

```

FEATURES
source
Location/Qualifiers
1..734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1009197"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="Vector: pME18SF13"

```

ORIGIN	Query Match	Score	DB 1	Length	DB 2	Score	DB 3	Length	DB 4	Score	DB 5	Length	DB 6	Score	DB 7	Length	DB 8	Score	DB 9	Length	DB 10	Score	DB 11	Length	DB 12	Score	DB 13	Length	DB 14	Score	DB 15	Length	DB 16	Score	DB 17	Length	DB 18	Score	DB 19	Length	DB 20	Score	DB 21	Length	DB 22	Score	DB 23	Length	DB 24	Score	DB 25	Length	DB 26	Score	DB 27	Length	DB 28	Score	DB 29	Length	DB 30	Score	DB 31	Length	DB 32	Score	DB 33	Length	DB 34	Score	DB 35	Length	DB 36	Score	DB 37	Length	DB 38	Score	DB 39	Length	DB 40	Score	DB 41	Length	DB 42	Score	DB 43	Length	DB 44	Score	DB 45	Length	DB 46	Score	DB 47	Length	DB 48	Score	DB 49	Length	DB 50	Score	DB 51	Length	DB 52	Score	DB 53	Length	DB 54	Score	DB 55	Length	DB 56	Score	DB 57	Length	DB 58	Score	DB 59	Length	DB 60	Score	DB 61	Length	DB 62	Score	DB 63	Length	DB 64	Score	DB 65	Length	DB 66	Score	DB 67	Length	DB 68	Score	DB 69	Length	DB 70	Score	DB 71	Length	DB 72	Score	DB 73	Length	DB 74	Score	DB 75	Length	DB 76	Score	DB 77	Length	DB 78	Score	DB 79	Length	DB 80	Score	DB 81	Length	DB 82	Score	DB 83	Length	DB 84	Score	DB 85	Length	DB 86	Score	DB 87	Length	DB 88	Score	DB 89	Length	DB 90	Score	DB 91	Length	DB 92	Score	DB 93	Length	DB 94	Score	DB 95	Length	DB 96	Score	DB 97	Length	DB 98	Score	DB 99	Length	DB 100	Score	DB 101	Length	DB 102	Score	DB 103	Length	DB 104	Score	DB 105	Length	DB 106	Score	DB 107	Length	DB 108	Score	DB 109	Length	DB 110	Score	DB 111	Length	DB 112	Score	DB 113	Length	DB 114	Score	DB 115	Length	DB 116	Score	DB 117	Length	DB 118	Score	DB 119	Length	DB 120	Score	DB 121	Length	DB 122	Score	DB 123	Length	DB 124	Score	DB 125	Length	DB 126	Score	DB 127	Length	DB 128	Score	DB 129	Length	DB 130	Score	DB 131	Length	DB 132	Score	DB 133	Length	DB 134	Score	DB 135	Length	DB 136	Score	DB 137	Length	DB 138	Score	DB 139	Length	DB 140	Score	DB 141	Length	DB 142	Score	DB 143	Length	DB 144	Score	DB 145	Length	DB 146	Score	DB 147	Length	DB 148	Score	DB 149	Length	DB 150	Score	DB 151	Length	DB 152	Score	DB 153	Length	DB 154	Score	DB 155	Length	DB 156	Score	DB 157	Length	DB 158	Score	DB 159	Length	DB 160	Score	DB 161	Length	DB 162	Score	DB 163	Length	DB 164	Score	DB 165	Length	DB 166	Score	DB 167	Length	DB 168	Score	DB 169	Length	DB 170	Score	DB 171	Length	DB 172	Score	DB 173	Length	DB 174	Score	DB 175	Length	DB 176	Score	DB 177	Length	DB 178	Score	DB 179	Length	DB 180	Score	DB 181	Length	DB 182	Score	DB 183	Length	DB 184	Score	DB 185	Length	DB 186	Score	DB 187	Length	DB 188	Score	DB 189	Length	DB 190	Score	DB 191	Length	DB 192	Score	DB 193	Length	DB 194	Score	DB 195	Length	DB 196	Score	DB 197	Length	DB 198	Score	DB 199	Length	DB 200	Score	DB 201	Length	DB 202	Score	DB 203	Length	DB 204	Score	DB 205	Length	DB 206	Score	DB 207	Length	DB 208	Score	DB 209	Length	DB 210	Score	DB 211	Length	DB 212	Score	DB 213	Length	DB 214	Score	DB 215	Length	DB 216	Score	DB 217	Length	DB 218	Score	DB 219	Length	DB 220	Score	DB 221	Length	DB 222	Score	DB 223	Length	DB 224	Score	DB 225	Length	DB 226	Score	DB 227	Length	DB 228	Score	DB 229	Length	DB 230	Score	DB 231	Length	DB 232	Score	DB 233	Length	DB 234	Score	DB 235	Length	DB 236	Score	DB 237	Length	DB 238	Score	DB 239	Length	DB 240	Score	DB 241	Length	DB 242	Score	DB 243	Length	DB 244	Score	DB 245	Length	DB 246	Score	DB 247	Length	DB 248	Score	DB 249	Length	DB 250	Score	DB 251	Length	DB 252	Score	DB 253	Length	DB 254	Score	DB 255	Length	DB 256	Score	DB 257	Length	DB 258	Score	DB 259	Length	DB 260	Score	DB 261	
--------	-------------	-------	------	--------	------	-------	------	--------	------	-------	------	--------	------	-------	------	--------	------	-------	------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	-------	-------	-------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--------	--------	-------	--------	--

RESULT 22	ACCESSION	ORGANISM
BX463976	VERSION	
LOCUS	KEYWORDS	
DEFINITION	SOURCE	

EX463976 852 bp mRNA linear EST 05-MAY-2004  
EX463976 Homo sapiens FETAL BRAIN Homo sapiens CDNA clone  
CS0DF021YA01 5-PRIME. mRNA sequence.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 852)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 22, 2003 this sequence version replaced gi:31023592.

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 8104.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?8=CSIAF0062A01QP1&c=8104.r.

## FEATURES

Location/Qualifiers  
1. .852  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DF021YX01"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-strand cDNA was digested with Not I and  
cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
vector. Library was not normalized."

## ORIGIN

Query Match 39.1%; Score 498; DB 5; Length 852;  
Best Local Similarity 99.8%; Pred. No. 3e-256;  
Matches 548; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 317 CAGAATGAGCGCTGACAGTGTCTACAGCATTACTACCGGTACCTCGAAGTGTCCACCAT 376  
DB 1 CAGAATGAGCGCTGACAGTGTCTACAGCATTACTACCGGTACCTCGAAGTGTCCACCAT 60  
QY 377 CCCAGAGGTGCTGCCCTGACTGGCACAACCTGCATCCCAACATGACTTTGGCGAGTCTTT 436  
DB 61 CCCAGAGGTGCTGCCCTGACTGGCACAACCTGCATCCCAACATGACTTTGGCGAGTCTTT 120  
QY 437 TTGAGTGTCCAGTCTGTTGACTATGTGTACCGGCATTCTTCATGTCAGAGTGCC 496  
DB 121 TTGAGTGTCCAGTCTGTTGACTATGTGTACCGGCATTCTTCATGTCAGAGTGCC 180  
QY 497 ATCTTGTGTTAGCACTGTGCGCCCAAGCTCACATGTTGTCCAACTTCCGCGGGCCCTT 556  
DB 181 ATCTTGTGTTAGCACTGTGCGCCCAAGCTCACATGTTGTCCAACTTCCGCGGGCCCTT 240  
QY 557 TGGGATCCATTCCCACTGGGTATGGAGAAAGTGCTTAATCAGTACTTTTCCCTGTA 616  
DB 241 TGGGATCCATTCCCACTGGGTATGGAGAAAGTGCTTAATCAGTACTTTTCCCTGTA 300  
QY 617 ATATGCGTCTTCTGGATGTGAATAACTCTGCGCACACAGAAAGCAGACCATGAG 676  
DB 301 ATATGCGTCTTCTGGATGTGAATAACTCTGCGCACACAGAAAGCAGACCATGAG 360  
QY 677 AGCTCTGTGAGTTTAGGCTTATCTCTGTCGTCCTGCTGCTTCTCTTAATGCAAG 736  
DB 361 AGCTCTGTGAGTTTAGGCTTATCTCTGTCGTCCTGCTGCTTCTCTTAATGCAAG 420  
QY 737 GCTCTCTGGATGCTGTAATGCCCACTCTGATGATCAGCATAGTCCATTACAACCTTAC 796  
DB 421 GCTCTCTGGATGCTGTAATGCCCACTCTGATGATCAGCATAGTCCATTACAACCTTAC 480  
QY 797 AGGAGAGAGATATAGTTTTTCTTGCTACAGACATTAACTTCTCTGCTGCTGTGACTGGG 856

Db 481 AGGAGAGATATAGTTTTTCTTGCTACAGACATTAACTTCTCTGCTGCTGTGACTGGG 540  
QY 857 TGATGATGG 865  
Db 541 TGATGATGC 549

## RESULT 23

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BX390856  
BX390856 Homo sapiens HELA CELLS 969 bp mRNA linear EST 29-APR-2004  
CDNA clone CS0DK012Y120 5-PRIME, mRNA sequence.  
BX390856  
BX390856.2 GI:46875200  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 969)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 13, 2003 this sequence version replaced gi:30619239.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
8104.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?8=CS0BAG050ZE02\_CS04717\_1&c=8104.r.

## FEATURES

Location/Qualifiers  
1. 969  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK012Y120"  
/cell\_type="HELA CELLS"  
/cell\_line="HELA"  
/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 39.0%; Score 497; DB 5; Length 969;  
Best Local Similarity 100.0%; Pred. No. 2.5e-255;  
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 710 GCCCTGGTCTTCTCTGTAATGCGCAAGCTCTCTGATGCTGTAAATGCCCATCTCATGC 769  
DB 1 GCCCTGGTCTTCTCTGTAATGCGCAAGCTCTCTGATGCTGTAAATGCCCATCTCATGC 60  
QY 770 ATCAGCAATAGTCCATTACACCCCTACAGGAGAGATATAGTTTTTCTTGCTACAGACA 829  
DB 61 ATCAGCAATAGTCCATTACACCCCTACAGGAGAGATATAGTTTTTCTTGCTACAGACA 120  
QY 830 TTAATCTTCTGCTGCTGTTGACTGGGTGATGATCAGTCCCTGTTTGGCTTTCACCTCA 889  
DB 121 TTAATCTTCTGCTGCTGTTGACTGGGTGATGATGAGTCTGTTTGGCTTTCACCTCA 180  
QY 890 TGTAGTCTTAGAGAAACAGGAAAAATACGATGGTCACCGAGTCTTCTCGCAATCGTAC 949  
DB 181 TGTAGTCTTAGAGAAACAGGAAAAATACGATGGTCACCGAGTCTTCTCGCAATCGTAC 240  
QY 950 AGCTGATAGGAACACGCAAGCAAGCTGAAAAATTTTCTTACCGACTTGAGCTTAATGGTC 1009



```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="PLACE1003647"  
/tissue_type="placenta"  
/clone_lib="PLACE1"  
/note="Vector: pME18SFL3"
```

## ORIGIN

Query Match 38.8%; Score 494; DB 1; Length 757;  
Best Local Similarity 100.0%; Pred. No. 1e-253;  
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

317	CAGAAATGAGCGT	CAGACTGCTACAGCATTAACCTACCGGTACCTCGAAGTGCCACCAT	376
154	CAGAAATGAGCGT	CAGACTGCTACAGCATTAACCTACCGGTACCTCGAAGTGCCACCAT	213
377	CCCAGAGGGTGCT	GCCTGACTGGCCACAACTGCATCCAAACAATGACTTGGCGAGTCTTTT	436
214	CCCAGAGGGTGCT	GCCTGACTGGCCACAACTGCATCCAAACAATGACTTGGCGAGTCTTTT	273
437	TTGAGTGTCCAGT	CTGCTGTGTGATATGTGTACCGCCCAATCTTTCAATGTCCAGATGGCC	496
274	TTGAGTGTCCAGT	CTGCTGTGTGATATGTGTACCGCCCAATCTTTCAATGTCCAGATGGCC	333
497	ATCTTGTTTGTAG	CAACTGTCCGCCAAGCTTCACATGTGTCCAACTGCGCGGGGCCCTT	556
334	ATCTTGTTTGTAG	CAACTGTCCGCCAAGCTTCACATGTGTCCAACTGCGCGGGGCCCTT	393
557	TGGGATCCATT	CGCAACTTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTGA	616
394	TGGGATCCATT	CGCAACTTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTGA	453
617	AATATGCGTCTT	CTCGATGTGAAATAAATCTGCCACACAGAAAAAGCAGACCATGAAG	676
454	AATATGCGTCTT	CTCGATGTGAAATAAATCTGCCACACAGAAAAAGCAGACCATGAAG	513
677	AGCTCTGTGAGT	TTTAGGCTTATTCCTGTCCGTGCCCTCGGTGCTTCCGTAAATGGCAAG	736
514	AGCTCTGTGAGT	TTTAGGCTTATTCCTGTCCGTGCCCTCGGTGCTTCCGTAAATGGCAAG	573
737	GCCTCTGGATGCT	GTAAATGCCCCCATCTGTATGCATCAGCATAAAGTCCATTACAACCCCTAC	796
574	GCTCTCTGGATGCT	GTAAATGCCCCCATCTGTATGCATCAGCATAAAGTCCATTACAACCCCTAC	633
797	AGGGAGAGGATATA	810	
634	AGGGAGAGGATATA	647	

RESULT 26  
 LOCUS BQ067628  
 DEFINITION BQ067628 1157 bp mRNA linear EST 02-APR-2002  
 AGENCOURT\_6758845 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5755201  
 5', mRNA sequence.  
 ACCESSION BQ067628  
 VERSION BQ067628.1 GI:19896674  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1157)  
 NTH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-i@mail.nih.gov](mailto:cgabs-i@mail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found at: <http://www.ncbi.nlm.nih.gov/ncbi/branches/mgc/>

RESULT 27

RESULT 27	ACCESSION
AU146134	VERSION
LOCUS	KEYWORDS
DEFINITION	

AUI46134 562 bp mRNA linear EST 05-AUG-2002  
AUI46134 HEMBA1 clone HEMBA1006927 3', mRNA  
sequence.  
AUI46134  
AUI46134.1 GI:11007655  
EST.

Db 481 ACTCTG 486

RESULT 28  
 BG435333 767 bp mRNA linear EST 14-MAR-2001  
 LOCUS 602508085F1 NIH\_MGC\_79 Homo sapiens cDNA clone IMAGE:4605404 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG435333  
 VERSION BG435333  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 767)  
 AUTHORS NTH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact Robert Strausberg, Ph.D.  
 Email: rga@biml.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LHC1342 row: i column: 21  
 High quality sequence stop: 695.  
 Location/Qualifiers  
 1. 767  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4605404"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 79"  
 /notes="Organ: placenta; Vector: pDNR-LIB (Clontech);  
 Site 1: SfII (ggcgcgcggcc); Site 2: SfiI  
 (ggcattatggcc); 5' and 3' adaptors were used in cloning  
 as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'  
 and 3' adaptor sequence:  
 5'-ATTCTAGAGCCGCGCGGCACATG-DT(30)BN-3' (where B = A,  
 C, G or N = A, C, G or T). Average insert size 1.3  
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH\_MGC Library."

FEATURES  
 source  
 1. 767  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4605404"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 79"  
 /notes="Organ: placenta; Vector: pDNR-LIB (Clontech);  
 Site 1: SfII (ggcgcgcggcc); Site 2: SfiI  
 (ggcattatggcc); 5' and 3' adaptors were used in cloning  
 as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'  
 and 3' adaptor sequence:  
 5'-ATTCTAGAGCCGCGCGGCACATG-DT(30)BN-3' (where B = A,  
 C, G or N = A, C, G or T). Average insert size 1.3  
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 38.1%; Score 486; DB 4; Length 767;  
 Best Local Similarity 99.8%; Pred. No. 2.1e-249  
 Matches 536; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 687 GTTTAGGCTTATTCCTGTCGTCGCGCCCTGGTGTCTCTCTGTAAATGGCAAGGCTCTCTGGA 746  
 Db 1 GTTTAGGCTTATTCCTGTCGTCGCGCCCTGGTGTCTCTCTGTAAATGGCAAGGCTCTCTGGA 60  
 Qy 747 TGTGTAATGCCCATCTGATGCAATCAGCAATAGTCATTCACGCTCAGGGAGAGGA 806  
 Db 61 TGTGTAATGCCCATCTGATGCAATCAGCAATAGTCATTCACGCTCAGGGAGAGGA 120  
 Qy 807 TATAGTTTTCTTGTCTACAGACATTAATCTTCTGTGTCTGTGTGACCTGGTGATGATGCA 866  
 Db 121 TATAGTTTTCTTGTCTACAGACATTAATCTTCTGTGTCTGTGTGACCTGGTGATGATGCA 180  
 Qy 867 GTCTGTTTTGGCTTTCACTTCATGTTAGTCTTTAGAGAACACAGAAAAATACCATGGTCA 926  
 Db 181 GTCTGTTTTGGCTTTCACTTCATGTTAGTCTTTAGAGAACACAGAAAAATACCATGGTCA 240  
 Qy 927 CCAGCAGTCTTCGCAATCGTACAGCTGATAGGAAACACGCAACAGCTGAAATTTTGC 986

Db 241 CCAGCAGTCTTCCGCAATCGTACAGCTGATAGGAAACAGCAGCAAGCTGAAAATTTTGC 300  
QY 987 TTACCACTGAGCTAAATGGTCATAGGCGACGATTTGACTTGGGAAGCGACTCCTCGATC 1046  
Db 301 TTACCACTGAGCTAAATGGTCATAGGCGACGATTTGACTTGGGAAGCGACTCCTCGATC 360  
QY 1047 TATTCAATGAAGGAATTCACACAGCCATTAAGATAGGCTGTCTAGTCTTTGACACCCAG 1106  
Db 361 TATTCAATGAAGGAATTCACACAGCCATTAAGATAGGCTGTCTAGTCTTTGACACCCAG 420  
QY 1107 CATTTGCACAGCTTTTTCAGAAAATGGCAATTTAGGCATCAATGTAACCTATTTCCATGTG 1166  
Db 421 CATTTGCACAGCTTTTTCAGAAAATGGCAATTTAGGCATCAATGTAACCTATTTCCATGTG 480  
QY 1167 TTGAATGGCAATCAACATTTTCGCCAGTGTTTAAACCTTCAGTTTCACAGAAA 1223  
Db 481 TCGAATGGCAATCAACATTTTCGCCAGTGTTTAAACCTTCAGTTTCACAGAAA 537

RESULT 29  
BE502849  
LOCUS  
DEFINITION  
h280b11.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3214269 3'  
similar to TR:043269 O43269 HSIAT1. ; mRNA sequence.  
ACCESSION  
BE502849  
VERSION  
BE502849.1 GI:9705257  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 647)  
AUTHORS  
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christophet Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lannon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
info@image.llnl.gov  
Seq primer: -40UP from Gibco  
High quality sequence stop: 448.  
Location/Qualifiers  
1..647  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3214269"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Lu24"  
/note="Organ: lung; Vector: pTV73D-Pac (Pharmacia) with a  
modified polylinker; Plasmid DNA from the normalized  
library NCI CGAP Lu5 was prepared, and ss circles were  
made in vitro. Following HAP purification, this DNA was  
used as tracer in a subtractive hybridization reaction.  
The driver was PCR-amplified cDNAs from a pool of 5,000  
clones made from the same library (clones  
1414920-1417991 and 1520904-1522439). Subtraction by Bento  
Soares and M. Fatima Bonaldo."

FEATURES  
source

Query Match 38.1%; Score 485; DB 2; Length 647;  
Best Local Similarity 99.8%; Pred. No. 7e-249;  
Matches 535; Conservative 0; Mismatches 1; Indels 0 Gaps 0;  
QY 521 CAAAGCTCAGTGTGTTCCAACTTGCAGGCGCCCTTTGGATCCATTCGCAACTTGGCTA 580

## ORIGIN

Query Match 38.1%; Score 485; DB 2; Length 647;  
Best Local Similarity 99.8%; Pred. No. 7e-249;  
Matches 535; Conservative 0; Mismatches 1; Indels 0 Gaps 0;  
QY 521 CAAAGCTCAGTGTGTTCCAACTTGCAGGCGCCCTTTGGATCCATTCGCAACTTGGCTA 580

Db 1 CAAAGCTCAGTGTGTTCCAACTTGCAGGCGCCCTTTGGATCCATTCGCAACTTGGCTA 60  
QY 581 TGGAGAAAGTGGCTAAATCAGTACTTTTCCCTGTAAATATGCTCTTCTGGATGTGAAA 640  
Db 61 TGGAGAAAGTGGCTAAATCAGTACTTTTCCCTGTAAATATGCTCTTCTGGATGTGAAA 120  
QY 641 TAACTCTGCCACACACAGAAAAGAGCAGACATGAAGAGCTCTGTGAGTTAGGCTTATT 700  
Db 121 TAACTCTGCCACACACAGAAAAGAGCAGACATGAAGAGCTCTGTGAGTTAGGCTTATT 180  
QY 701 CTTGTCCTGCGCTGCTGCTCTCTGTAATGCAAGGCTCTCTGGAATGCTGTAAATGCC 760  
Db 181 CTTGTCCTGCGCTGCTGCTCTCTGTAATGCAAGGCTCTCTGGAATGCTGTAAATGCC 240  
QY 761 ATCTGATGATCAGCATTAAGTCCATTCAACCCCTACAGGAGAGGATATAGTTTCTTG 820  
Db 241 ATCTGATGATCAGCATTAAGTCCATTCAACCCCTACAGGAGAGGATATAGTTTCTTG 300  
QY 821 CTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880  
Db 301 CTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
QY 881 TTCACTTCAATGTTAGTCTTTAGAGAAACAGGAAAATATGATGCTCACCAGAGTTCTTCG 940  
Db 361 TTCACTTCAATGTTAGTCTTTAGAGAAACAGGAAAATATGATGCTCACCAGAGTTCTTCG 420  
QY 941 CAATCGTACAGCTGATAGGAAACAGCAAGCAAGCTGAAAATTTTGTACCCGACTTGAGC 1000  
Db 421 CAATCGTACAGCTGATAGGAAACAGCAAGCAAGCTGAAAATTTTGTACCCGACTTGAGC 480  
QY 1001 TAAATGGTTCATAGGCGACGATTTGACTTGGGAAGCGACTCCTCGATCTATTTCATGAA 1056  
Db 481 TAAATGGTTCATAGGCGACGATTTGACTTGGGAAGCGACTCCTCGATCTATTTCATGAA 536

RESULT 30  
BU191771  
LOCUS  
DEFINITION  
AGENCOURT 7779845 NIH\_MGC\_70 Homo sapiens cDNA clone IMAGE:6019472  
5', mRNA sequence.  
ACCESSION  
BU191771  
VERSION  
BU191771.1 GI:22705755  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 976)  
AUTHORS  
NIH-MGC http://mgc.nci.nih.gov/.  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13222 row: a column: 09  
High quality sequence stop: 446.  
Location/Qualifiers  
1..976  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6019472"  
/tissue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_70"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site\_1: NotI;



Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."

ORIGIN  
Query Match 38.1%; Score 485; DB 5; Length 976;  
Best Local Similarity 100.0%; Pred. No. 7.3e-249;  
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 641 TAACTCTGCCACACAGAAAAGAGACCATGAGAGCTCTGTAGTTAGGCGCTTATT 700  
Db 1 TAACTCTGCCACACAGAAAAGAGACCATGAGAGCTCTGTAGTTAGGCGCTTATT 60  
QY 701 CTTGTCGCGGCCCTGCTCTCTGTAATGCAAGGCTCTCTGATGCTGTAAATGCCCC 760  
Db 61 CTTGTCGCGGCCCTGCTCTCTGTAATGCAAGGCTCTCTGATGCTGTAAATGCCCC 120  
QY 761 ATCTGATGATCAGCATAGTCCATTACAACTTACAGGAGAGATATAGTTTCTTG 820  
Db 121 ATCTGATGATCAGCATAGTCCATTACAACTTACAGGAGAGATATAGTTTCTTG 180  
QY 821 CTACAGACATTAATCTCTGCTGTGCTGTGATGATGATGATGATGATGATGATGAT 880  
Db 181 CTACAGACATTAATCTCTGCTGTGCTGTGATGATGATGATGATGATGATGATGAT 240  
QY 881 TTTCACTTCACTGTTAGTCTTGTAGAGAACAGAGAAATACGATGCTCAACAGAGTCTTCG 940  
Db 241 TTTCACTTCACTGTTAGTCTTGTAGAGAACAGAGAAATACGATGCTCAACAGAGTCTTCG 300  
QY 941 CAATCGTACAGCTGATAGGAAACAGCAAGCTGAAATTTTGTACCGACTTGAGC 1000  
Db 301 CAATCGTACAGCTGATAGGAAACAGCAAGCTGAAATTTTGTACCGACTTGAGC 360  
QY 1001 TAAATGCTATAGGACAGCAATGACTTGGAGAGCACTCTCGATCTATTATGAGGAA 1060  
Db 361 TAAATGCTATAGGACAGCAATGACTTGGAGAGCACTCTCGATCTATTATGAGGAA 420  
QY 1061 TTGCAACAGCCATTATGATAGGAGTCTAGTCTTTGACACCAAGCAATGACAGCTTT 1120  
Db 421 TTGCAACAGCCATTATGATAGGAGTCTAGTCTTTGACACCAAGCAATGACAGCTTT 480  
QY 1121 TTGCA 1125  
Db 481 TTGCA 485

RESULT 31  
CF551876 890 bp mRNA linear EST 22-SEP-2003  
LOCUS  
DEFINITION  
AGENCOURT 15595245 NIH\_MGC\_183 Homo sapiens cDNA clone  
IMAGE:30530108 5', mRNA sequence.

CF551876  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 890)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [cgapsb-remail.nih.gov](mailto:cgapsb-remail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>  
Plate: NDAM618 row: 1 column: 21  
High quality sequence start: 6  
High quality sequence stop: 590.

FEATURES  
Source  
Location/Qualifiers  
1..890  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30530108"  
/lab\_host="DH10B-Tona (T1 and T5 phage resistant)"  
/clone\_lib="NIH\_MGC\_183"  
/note="Organ: Pooled muscle (cardiac and skeletal);  
Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2:  
NotI; Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.7. Library was constructed by Invitrogen."

ORIGIN  
Query Match 37.4%; Score 477; DB 7; Length 890;  
Best Local Similarity 100.0%; Pred. No. 1.5e-244;  
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 599 CAGTACTTTTCCCTGTAATATGCTCTCTGATGTAATTAATTAATTAATTAATTAAT 658  
Db 168 CAGTACTTTTCCCTGTAATATGCTCTCTGATGTAATTAATTAATTAATTAATTAAT 227  
QY 659 AAAAAGCAGACCATGAAGAGCTCTGTGATTTAGGCTTATTCTCTGCTCCCTGGTG 718  
Db 228 AAAAAGCAGACCATGAAGAGCTCTGTGATTTAGGCTTATTCTCTGCTCCCTGGTG 287  
QY 719 CTTCTGTAATGGAAGGCTCTCTGATGCTTAATGCCCATCTGATGATGATGATGATGAT 778  
Db 288 CTTCTGTAATGGAAGGCTCTCTGATGCTTAATGCCCATCTGATGATGATGATGATGAT 347  
QY 779 AGTCATTACAACTTACAGGAGAGGATATAGTTTCTGCTACAGACATTAATCTTC 838  
Db 348 AGTCATTACAACTTACAGGAGAGGATATAGTTTCTGCTACAGACATTAATCTTC 407  
QY 839 CTGCTGCTGTTGACTGGGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 898  
Db 408 CTGCTGCTGTTGACTGGGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 467  
QY 899 TAGAAGAAACAGAAAATACGATGCTCAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 958  
Db 468 TAGAAGAAACAGAAAATACGATGCTCAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 527  
QY 959 GAACAGCAAGCAAGCTGAAATTTGCTTACCGACTGAGCTAAATGCTCATAGCGCAC 1018  
Db 528 GAACAGCAAGCAAGCTGAAATTTGCTTACCGACTGAGCTAAATGCTCATAGCGCAC 587  
QY 1019 GATTGACTTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATTGCAAGCCATTA 1075  
Db 588 GATTGACTTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATTGCAAGCCATTA 644

RESULT 32  
AI803300 515 bp mRNA linear EST 13-DEC-1999  
LOCUS  
DEFINITION  
tc17d08.x1 Soares NHPu S1 Homo sapiens cDNA clone IMAGE:2064111  
3' similar to TR:043269 O43269 HSA1H1.1; mRNA sequence.  
AI803300  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 515)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL





```
Db      361 TTGCACAGCTTTTTCAGAGAAATGCCAATTAGGCATCAATGTAACATATTCCATGTGTC 420
Qy      1169 GAAATGGCAATCAAAACATTTTCGCCAGTGTTTAAACCTTCAGTTTCACAGAAATAAG 1228
Db      421 GAAATGGCAATCAAAACATTTTTCGCCAGTGTTTAAACCTTCAGTTTCACAGAAATAAG 480
Qy      1229 GCACCCATCTGTCTGCCACCTTAAACCTTTTCGGTAGTGGAGC 1274
Db      481 GCACCCATCTGTCTGCCACCTTAAACCTTTTCGGTAGTGGAGC 526

RESULT 34
LOCUS   BX384656
DEFINITION BX384656 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
ACCESSION BX384656
VERSION   BX384656
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 785)
AUTHORS   Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   On May 8, 2003 this sequence version replaced gi:30447286.
          Contact: Genoscope
          Genoscope - Centre National de Sequencage
          2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
          Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
          1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
          end enriched, double-strand cDNA was digested with Not I and cloned
          into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
          was normalized. Library was constructed by Life Technologies, a
          division of Invitrogen. This sequence belongs to sequence cluster
          8104.r
          For more information about this cluster, see
          http://www.genoscope.cns.fr/cdna?8=CSODK012BE100Plc=8104.r.

FEATURES
        source
            1..785
                Location/Qualifiers
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="CSODK012Y120"
                    /cell_line="HELA"
                    /clone_lib="HELA CELLS COT 25-NORMALIZED"
                    /notes="1st strand cDNA was primed with a NotI-oligo(dT)
                    primer. Five prime end enriched, double-strand cDNA was
                    digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      37.3%; Score 475; DB 5; Length 785;
Best Local Similarity 99.8%; Pred. No. 1.7e-243;
Matches 525; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      317 CAGAAATGAGCCCTGACAGCTGTACAGCTACTACCGGTACCTCGAAGTGTCCACCAT 376
Db      1 CAGAAATGAGCCCTGACAGCTGTACAGCTACTACCGGTACCTCGAAGTGTCCACCAT 60
Qy      377 CCAGAGGGTGGCTGCGCTGACTGGGACACTGCATCCCAATGACTTGGCCAGTCTTT 436
Db      61 CCAGAGGGTGGCTGCGCTGACTGGGACACTGCATCCCAATGACTTGGCCAGTCTTT 120
Qy      437 TTGAGTGTCCAGTCTGCTTTGACTATGTGTATCGGCCCATTTCTCAATGTCAGAGTGGCC 496
Db      121 TTGAGTGTCCAGTCTGCTTTGACTATGTGTATCGGCCCATTTCTCAATGTCAGAGTGGCC 180
Qy      497 ATCTTGTTTGTAGCACTGTGCGCCCAAGCTCACATGTTGTCCAACTTGGCGGGGCCCTT 556
```

```
Db      181 ATCTTGTTTGTAGCACTGTGCGCCCAAGCTCACATGTTGTTCAACTTGGCGGGGCCCTT 240
Qy      557 TGGGATCATTTCGCAACTTGGCTATGAGAGAAAGTGGCTTAATTCAGTACTTTTCCCTGTA 616
Db      241 TGGGATCATTTCGCAACTTGGCTATGAGAGAAAGTGGCTTAATTCAGTACTTTTCCCTGTA 300
Qy      617 AATATGGCTCTTCTGGATGTGAATAACTCTGCACACACAGAAAAAGACGACCATGAAG 676
Db      301 AATATGGCTCTTCTGGATGTGAATAACTCTGCACACACAGAAAAAGACGACCATGAAG 360
Qy      677 AGCTCTGTGAGTTTAGGCTTATTCCTGTGCGTGGCTTCTCTGTAATATGGCAAG 736
Db      361 AGCTCTGTGAGTTTAGGCTTATTCCTGTGCGTGGCTTCTCTGTAATATGGCAAG 420
Qy      737 GCTCTCTGATGCTGTAAATGCCCCCATCTGATGTCATCAGCATAAGTCCATTACAACCTTAC 796
Db      421 GCTCTCTGATGCTGTAAATGCCCCCATCTGATGTCATCAGCATAAGTCCATTACAACCTTAC 480
Qy      797 AGGAGAGGATATAGTTTTTCTTGTCTACAGACATTAATCTTCCTGG 842
Db      481 AGGAGAGGATATAGTTTTTCTTGTCTACAGACATTAATCTTCCTGG 526

RESULT 35
LOCUS   BE503207
DEFINITION BE503207 Homo sapiens CDNA clone IMAGE:3214717 3'
          similar to TR:043269 O43269 HSA1H1.1, mRNA sequence.
ACCESSION BE503207
VERSION   BE503207
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 647)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
          Unpublished (1997)
          Contact: RGG@nci.nih.gov
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
          Emmert-Buck, M.D., Ph.D.
          CDNA Library Preparation: M. Bento Soares, Ph.D.
          CDNA Library Arrayed by: Greg Lennon, Ph.D.
          DNA Sequencing by: Washington University Genome Sequencing Center
          Clone distribution: NCI-CGAP clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL, send email to:
          info@image.llnl.gov
          Seq primer: -40UP from Gibco
          High quality sequence stop: 456.
          Location/Qualifiers
              1..647
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:3214717"
                  /tissue_type="carcinoid"
                  /lab_host="DH10B"
                  /clone_lib="NCI CGAP Lu24"
                  /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
                  modified polylinker; Plasmid DNA from the normalized
                  library NCI CGAP Lu5 was prepared, and ss circles were
                  made in vitro. Following HAP purification, this DNA was
                  used as tracer in a subtractive hybridization reaction.
                  The driver was PCR-amplified cDNAs from a pool of 5,000
                  clones made from the same library (cloneIDs
                  141920-1417991 and 152094-1522439). Subtraction by Bento
                  Soares and M. Fatima Bonaldo."

ORIGIN
Query Match      37.2%; Score 474; DB 2; Length 647;
```

```
Best Local Similarity 99.7%; Pred. No. 5.8e-243;
Matches 574; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 521 CAAAGCTCACATGTTGTCCAACTTGGCGGGCCCTTTGGGATCCATTCGCAACTTGGCTA 580
    |||||
Db 1 CAAAGCTCACATGTTGTCCAACTTGGCGGGCCCTTTGGGATCCATTCGCAACTTGGCTA 60

QY 581 TGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATATGCTCTTCGATGTGAAA 640
    |||||
Db 61 TGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATATGCTCTTCGATGTGAAA 120

QY 641 TAACTCTGCCACACACAGAAAAGCAGACCATCAAGAGCTCTGTAGTTTGGSCCTTATT 700
    |||||
Db 121 TAACTCTGCCACACACAGAAAAGCAGACCATCAAGAGCTCTGTAGTTTGGSCCTTATT 180

QY 701 CTTGTCCGTGCCCTGGTCTTCTGTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCC 760
    |||||
Db 181 CTTGTCCGTGCCCTGGTCTTCTGTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCC 240

QY 761 ATCTGATGATCAGCATATATTCATTAACACCTTACAGGGAGAGGATATAGTTTCTTGG 820
    |||||
Db 241 ATCTGATGATCAGCATATATTCATTAACACCTTACAGGGAGAGGATATAGTTTCTTGG 300

QY 821 CTACAGACATTAATCTCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
    |||||
Db 301 CTACAGACATTAATCTCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

QY 881 TTCACCTTCACTGTTAGTCTTTAGAGAAACAGAAAATACAGATGCTCACCAGCAGTCTTCG 940
    |||||
Db 361 TTCACCTTCACTGTTAGTCTTTAGAGAAACAGAAAATACAGATGCTCACCAGCAGTCTTCG 420

QY 941 CAATCTACAGCTGATAGGAACACGCAAGCACTGAAATTTTGTCTTACCGACTTGGAGC 1000
    |||||
Db 421 CAATCTACAGCTGATAGGAACACGCAAGCACTGAAATTTTGTCTTACCGACTTGGAGC 480

QY 1001 TAAATGGTCATAGGCAAGTACCTTGGGAAGGCTCTCTCGATCTATTTCATGAGGAA 1060
    |||||
Db 481 TAAATGGTCATAGGCAAGTACCTTGGGAAGGCTCTCTCGATCTATTTCATGAGGAA 540

QY 1061 TTGCAACAGCCATTATGAATAGGCACTGTCTAGTCT 1096
    |||||
Db 541 TTGCAACAGCCATTATGAATAGGCACTGTCTAGTCT 576

RESULT 36
BI836179
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

## FEATURES

```
source 1.800
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5224755"
/lab_host="DH10B"
/clone_lib="NIH_MGC_120"
/notes="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 37.2%; Score 474; DB 4; Length 800;
Best Local Similarity 100.0%; Pred. No. 5.9e-243;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 CAGAAATGAGCCCTCAGACTGTCTACAGCATTACTTACCGGTACTCTGAACTGTCCACCAT 376
    |||||
Db 111 CAGAAATGAGCCCTCAGACTGTCTACAGCATTACTTACCGGTACTCTGAACTGTCCACCAT 170

QY 377 CCAGAGGGTGGCTGCCCTGACTGGCACAACCTGCATCCAACTGCTTGGCGAGTCTTT 436
    |||||
Db 171 CCAGAGGGTGGCTGCCCTGACTGGCACAACCTGCATCCAACTGCTTGGCGAGTCTTT 230

QY 437 TTGAGTGTCACTGCTCTTGTGATGTATTACCGCCCAATTTCTCAATGTCAAGTGGCC 496
    |||||
Db 231 TTGAGTGTCACTGCTCTTGTGATGTATTACCGCCCAATTTCTCAATGTCAAGTGGCC 290

QY 497 ATCTGTTTGTAGCACTCTGCCCAAGCTCACATGTTGTGCAATGTCGCGGGCCCTT 556
    |||||
Db 291 ATCTGTTTGTAGCACTCTGCCCAAGCTCACATGTTGTGCAATGTCGCGGGCCCTT 350

QY 557 TGGGATCCATTGCACTTGGCTATGAGAAAGTGGCTAATTCAGTACTTTCCCTGTGA 616
    |||||
Db 351 TGGGATCCATTGCACTTGGCTATGAGAAAGTGGCTAATTCAGTACTTTCCCTGTGA 410

QY 617 AATATGCGTCTTCTGATGTGAAATAACTCTGCGCACACAGAAAAGCAGACCATGAAG 676
    |||||
Db 411 AATATGCGTCTTCTGATGTGAAATAACTCTGCGCACACAGAAAAGCAGACCATGAAG 470

QY 677 AGCTGTGAGTTAGGCTTATTCCTGTCGCGCCCTGCTGCTCTGTAATGCAAG 736
    |||||
Db 471 AGCTGTGAGTTAGGCTTATTCCTGTCGCGCCCTGCTGCTCTGTAATGCAAG 530

QY 737 GCTCTCTGGATGCTGTAATGCCCCATCTGATGATCAGCATAGTCCATTACAA 790
    |||||
Db 531 GCTCTCTGGATGCTGTAATGCCCCATCTGATGATCAGCATAGTCCATTACAA 584

RESULT 37
BX459114
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-Oligo (dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 8104.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?e=CS0DE012DG11QPI&c=8104.r.

FEATURES  
Location/Qualifiers  
1. 842  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE012YN22"  
/tissue\_type="PLACENTA"  
/clone\_lib="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-Oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

ORIGIN  
Query Match 37.1%; Score 473; DB 5; Length 842;  
Best Local Similarity 100.0%; Pred. No. 2.1e-242;  
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 434 TTTTTCAGTGTCAGTCTGCTTGTACTATGTGTACCGCCCATTTCTCAATGTCAGAGTG 493  
DB 114 TTTTTCAGTGTCAGTCTGCTTGTACTATGTGTACCGCCCATTTCTCAATGTCAGAGTG 173  
QY 494 GCCATCTGTTGTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCAACTTCGCGGGGCC 553  
DB 174 GCCATCTGTTGTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCAACTTCGCGGGGCC 233  
QY 554 CTTTGGATCCATTCGCACTGGCTATGGAGAAAGTGTCTTTCAGTACTTTTCCCT 613  
DB 234 CTTTGGATCCATTCGCACTGGCTATGGAGAAAGTGTCTTTCAGTACTTTTCCCT 293  
QY 614 GTAAATATGCGTCTTCTGGATGTGAATAACTCTGCCACACACAGAAAAGCAGACCATG 673  
DB 294 GTAAATATGCGTCTTCTGGATGTGAATAACTCTGCCACACACAGAAAAGCAGACCATG 353  
QY 674 AAGAGCTCTGTGATTTAGCCCTTATTCCTGTCGCGCCCTGGTGTCTTCTGTAATGGC 733  
DB 354 AAGAGCTCTGTGATTTAGCCCTTATTCCTGTCGCGCCCTGGTGTCTTCTGTAATGGC 413  
QY 734 AAGGCTCTGCGATGCTGTAATGCCCATCTGATGCATCAGATAAGTCCATTACACCC 793  
DB 414 AAGGCTCTGCGATGCTGTAATGCCCATCTGATGCATCAGATAAGTCCATTACACCC 473  
QY 794 TACAGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGGTGTGTTGAT 853  
DB 474 TACAGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGGTGTGTTGAT 533  
QY 854 GGGTATGATGCGATCTCTGTTTGGCTTTCACATTCATGTTAGTCTTAGAGAA 906  
DB 534 GGGTATGATGCGATCTCTGTTTGGCTTTCACATTCATGTTAGTCTTAGAGAA 586

RESULT 38  
BF573693  
LOCUS 894 bp mRNA linear EST 12-DEC-2000  
DEFINITION 602077289F2 NIH\_MGC\_62 Homo sapiens cDNA clone IMAGE:4251575 5',  
mRNA sequence.  
ACCESSION BF573693  
KEYWORDS BF573693.1 GI:11647405  
SOURCE EST.  
ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 894)  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: rgs@nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L10M1075 row: b column: 24  
High quality sequence stop: 562.  
Location/Qualifiers  
1. 894  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE.4251575"  
/tissue\_type="melanotic melanoma, high MDR"  
/lab\_host="BHI0B (T1 phase-resistant)"  
/clone\_lib="NIH\_MGC\_62"  
/note="Organ: skin; Vector: pDNR-LIB (Clontech); Site: 1:  
SfiI (ggcccttcggcc); Site 2: SfiI (ggccattatggcc);  
Double-stranded cDNA was prepared from cell line RNA. 5',  
and 3' adaptors were used in cloning as follows: 5',  
adaptor sequence 5'-ATTCTAGAGCCGAGGCGGCGGACATG-dT(30)BN-3',  
sequence: 5'-ATTCTAGAGCCGAGGCGGCGGCGGACATG-dT(30)BN-3',  
(where B = A, C, G, or T). Average  
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."

FEATURES  
source  
Location/Qualifiers  
1. 894  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE.4251575"  
/tissue\_type="melanotic melanoma, high MDR"  
/lab\_host="BHI0B (T1 phase-resistant)"  
/clone\_lib="NIH\_MGC\_62"  
/note="Organ: skin; Vector: pDNR-LIB (Clontech); Site: 1:  
SfiI (ggcccttcggcc); Site 2: SfiI (ggccattatggcc);  
Double-stranded cDNA was prepared from cell line RNA. 5',  
and 3' adaptors were used in cloning as follows: 5',  
adaptor sequence 5'-ATTCTAGAGCCGAGGCGGCGGACATG-dT(30)BN-3',  
sequence: 5'-ATTCTAGAGCCGAGGCGGCGGCGGACATG-dT(30)BN-3',  
(where B = A, C, G, or T). Average  
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."

ORIGIN  
Query Match 37.1%; Score 473; DB 2; Length 894;  
Best Local Similarity 100.0%; Pred. No. 2.1e-242;  
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 504 TTGTAGCAACTGTGCGCCAAAGCTCACATGTTTCCAACTTCCCGGGCCCTTTGGATC 563  
DB 105 TTGTAGCAACTGTGCGCCAAAGCTCACATGTTTCCAACTTCCCGGGCCCTTTGGATC 164  
QY 564 CATTGCGCACTTGGCTATGGAGAAAGTGGCTAATTCATTAATTTCCCTGTAATATGC 623  
DB 165 CATTGCGCACTTGGCTATGGAGAAAGTGGCTAATTCATTAATTTCCCTGTAATATGC 224  
QY 624 GTCTTCTGGATGTAATAACTCTGCCACACACAGAAAAGCAGACCATGAAGAGCTCTG 683  
DB 225 GTCTTCTGGATGTAATAACTCTGCCACACACAGAAAAGCAGACCATGAAGAGCTCTG 284  
QY 684 TGAGTTAGGCTTATTCCTGTCGCGCCCTGCTTCTTAAATGCAAGGCTCTCT 743  
DB 285 TGAGTTAGGCTTATTCCTGTCGCGCCCTGCTTCTTAAATGCAAGGCTCTCT 344  
QY 744 GGATGCTGTAATGCCCATCTGATGCATCAGATAAGTCCATTACACCTACAGGAGA 803  
DB 345 GGATGCTGTAATGCCCATCTGATGCATCAGATAAGTCCATTACACCTACAGGAGA 404  
QY 804 GGATATAGTTTTTCTGCTACAGACATTAATCTTCTGCTGCTGTTGATGATGAT 863  
DB 405 GGATATAGTTTTTCTGCTACAGACATTAATCTTCTGCTGCTGTTGATGATGAT 464  
QY 864 GCAGTCTCTTTGGCTTTCACCTTCTAGTAGAGAAACAGAGAAATACCATGG 923  
DB 465 GCAGTCTCTTTGGCTTTCACCTTCTAGTAGAGAAACAGAGAAATACCATGG 524  
QY 924 TCACCAAGCAGTCTTTCGCAATCTGATAGTATAGAGAACACGCAAGCAGCTG 976

```
Db 525 TCACCAGCAGTCTTCGCAATCTGTACAGCTGATAGGACACGCAAGCAAGCTG 577

RESULT 39
BP227842
LOCUS
DEFINITION
  BP227842 Sugano cDNA library, dermoid cancer Homo sapiens CDNA
  clone DMC07313, mRNA sequence.
ACCESSION
  BP227842
VERSION
  BP227842.1 GI:52100747
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens (human)
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 583)
AUTHORS
  Suzuki.Y., Yamashita.R., Shirota.M., Sakakibara.Y., Chiba.J.,
  Mizushima-Sugano.J., Nakai.K. and Sugano.S.
TITLE
  Sequence comparison of human and mouse genes reveals a homologous
  block structure in the promoter regions
JOURNAL
  Genome Res. 14 (9), 1711-1718 (2004)
COMMENT
  Contact: Yutaka Suzuki
  Department of Virology
  Institute of Medical Science, University of Tokyo
  4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
  Email: ysuzuki@im.s.u-tokyo.ac.jp.

FEATURES
  source
    1..583
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="DMC07313"
    /note="dermoid cancer"

ORIGIN
  Query Match 36.9%; Score 470; DB 5; Length 583;
  Best Local Similarity 100.0%; Pred. No. 8.2e-241;
  Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 317 CAGAAATGAGCGCTGACAGCTGTCTAGAGCAATCTACCGGTACCTCGAAGTGTCCACCAT 376
  Db 114 CAGAAATGAGCGCTGACAGCTGTCTAGAGCAATCTACCGGTACCTCGAAGTGTCCACCAT 173

  QY 377 CCAGAGGGTGGCTGCGCTGACTGGCACTGCAATCCCAATGACTTGGCGAGTCTTT 436
  Db 174 CCAGAGGGTGGCTGCGCTGACTGGCACTGCAATCCCAATGACTTGGCGAGTCTTT 233

  QY 437 TTGAGTGTCCAGTCTGCTTTGACTATGTGTACCGGCCAATCTTCAATGTGAGTGGCC 496
  Db 234 TTGAGTGTCCAGTCTGCTTTGACTATGTGTACCGGCCAATCTTCAATGTGAGTGGCC 293

  QY 497 ATCTGTTTGTAGCAACTGTGCGCCAAAGCTGACATGTTGTCCAACTTCGCGGGGCCCTT 556
  Db 294 ATCTGTTTGTAGCAACTGTGCGCCAAAGCTGACATGTTGTCCAACTTCGCGGGGCCCTT 353

  QY 557 TGGGATCCATTCGCAACTTGGCTATGAGAGAAAGTGGCTAAATTCAGTACTTTTCCCCTGTA 616
  Db 354 TGGGATCCATTCGCAACTTGGCTATGAGAGAAAGTGGCTAAATTCAGTACTTTTCCCCTGTA 413

  QY 617 AATATGCGTCTTCTGATGTGAAATTAATCTCTGCAACACAGAGAAAGCAGACCATGAAG 676
  Db 414 AATATGCGTCTTCTGATGTGAAATTAATCTCTGCAACACAGAGAAAGCAGACCATGAAG 473

  QY 677 AGCTGTGAGTTTAGGCTTATCTCTGTCGCGCCCTGCTCTTCTCTGTAATGGCAAG 736
  Db 474 AGCTGTGAGTTTAGGCTTATCTCTGTCGCGCCCTGCTCTTCTCTGTAATGGCAAG 533

  QY 737 GCTCTCTGATGCTGTAATGCCCCCATCTGATGCATCAGCATTAAGTCCATT 786
  Db 534 GCTCTCTGATGCTGTAATGCCCCCATCTGATGCATCAGCATTAAGTCCATT 583

  RESULT 40
  BX372197/c
  LOCUS
  DEFINITION
    BX372197 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
    cDNA clone CS0DK012YI20 3-PRIME, mRNA sequence.
  ACCESSION
    BX372197
  VERSION
    BX372197.2 GI:46573987
  SOURCE
    Homo sapiens (human)
  ORGANISM
    Homo sapiens (human)
  REFERENCE
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
    1 (bases 1 to 957)
  AUTHORS
    Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE
    Full-length cDNA libraries and normalization
  JOURNAL
    Unpublished (2001)
  COMMENT
    On May 8, 2003 this sequence version replaced gi:30450080.
    Contact: Genoscope
    Genoscope - Centre National de Sequencage
    2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
    Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
    1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
    end enriched, double-strand cDNA was digested with Not I and cloned
    into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
    was normalized. Library was constructed by Life Technologies, a
    division of Invitrogen. This sequence belongs to sequence cluster
    8104.r
    For more information about this cluster, see
    http://www.genoscope.cns.fr/cdna?S=CS0BAI041ZD02_CS03852_1&c=8104.r

  FEATURES
    source
      1..957
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DK012YI20"
      /cell_type="HELA CELLS COT 25-NORMALIZED"
      /cell_lines="HELA"
      /note="1st strand cDNA was primed with a NotI-oligo(dT)
      primer. Five prime end enriched, double-strand cDNA was
      digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

  ORIGIN
    Query Match 36.3%; Score 462; DB 5; Length 957;
    Best Local Similarity 100.0%; Pred. No. 1.7e-216;
    Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 813 TTTTCTTGCTACAGACATTAATCTTCTGGTGTGTGACTGGGTGATGATGAGTCTCG 872
    Db 610 TTTTCTTGCTACAGACATTAATCTTCTGGTGTGTGACTGGGTGATGATGAGTCTCG 551

    QY 873 TTTTGGCTTTCATCTTCATGTTAGTCTTAGAGAAACAGGAAATAGGTCACGCA 932
    Db 550 TTTTGGCTTTCATCTTCATGTTAGTCTTAGAGAAACAGGAAATAGGTCACGCA 491

    QY 933 GTTCTTTCGCAATCGTACAGCTGATAGGAAACAGCAAGCTGAAATTTTCTTACCG 992
    Db 490 GTTCTTTCGCAATCGTACAGCTGATAGGAAACAGCAAGCTGAAATTTTCTTACCG 431

    QY 993 ACTTGAGCTAAATGTCATAGCGCGAATTTGGAAGGAGTCTCTCGATCTATTCA 1052
    Db 430 ACTTGAGCTAAATGTCATAGCGCGAATTTGGAAGGAGTCTCTCGATCTATTCA 371

    QY 1053 TGAAGCAATTCGCAACAGCCATTAATGATGCACTGTCTAGTCTTTCACACCATTCG 1112
    Db 370 TGAAGCAATTCGCAACAGCCATTAATGATGCACTGTCTAGTCTTTCACACCATTCG 311

    QY 1113 ACAGCTTTTTCAGAAAAATGGCAATTTTAGGCATCAATGTAATTTCCATGTTGAAA 1172
    Db 310 ACAGCTTTTTCAGAAAAATGGCAATTTTAGGCATCAATGTAATTTCCATGTTGAAA 251
```

```

QY 1173 TGGCAATCAAACTTTCTGGCCAGTGTGTTAAATCTTCAGTTTCACAGAAATAAGGCAC 1232
Db 250 TGGCAATCAAACTTTCTGGCCAGTGTGTTAAATCTTCAGTTTCACAGAAATAAGGCAC 191
QY 1233 CCATCTGTCTGCGCAACTAAATCTTTTCGGTAGTGGGAAGC 1274
Db 190 CCATCTGTCTGCGCAACTAAATCTTTTCGGTAGTGGGAAGC 149

RESULT 41
AA917325
LOCUS on45b05.s1 NCI CGAP Co8 Homo sapiens cDNA clone IMAGE:1559601 3'
DEFINITION similar to TR:Q06984 Q06984 SEVEN IN ABSENITIA 1A ; mRNA sequence.
ACCESSION AA917325
VERSION AA917325
KEYWORDS A0917325.1 GI:3057215
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 513)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 1036 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 159.
Location/Qualifiers
FEATURES
source
1..513
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1559601"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Co8"
/notes="Organ: Colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaudo."
ORIGIN
Query Match 36.0%; Score 459; DB 1; Length 513;
Best Local Similarity 100.0%; Pred. No. 6.7e-235; Indels 0; Gaps 0;
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 TTTATTTTAAATATCTTTTAAATCCCTATTTTCTCTCTTTGCTCAGTAAATTTT 219
Db 1 TTTATTTTAAATATCTTTTAAATCCCTATTTTCTCTCTTTGCTCAGTAAATTTT 60

QY 220 GTATGAACCTTTAAAGGACTTATGGCATGTAACATTTATTAAGTAAGTCATGGTT 279
Db 61 GTATGAACCTTTAAAGGACTTATGGCATGTAACATTTATTAAGTAAGTCATGGTT 120

QY 280 ATAATATTTTCTCTCGCTCTCTTATGATTTATTTTACAGAAATGAGCCCTCAGACTGCT 339
Db 121 ATAATATTTTCTCTCGCTCTCTTATGATTTATTTTACAGAAATGAGCCCTCAGACTGCT 180

```

```

QY 340 ACAGCATTTACCTACCGGTACCTCGAAGTGTCCACATCCAGAGGGTGCCTGCCTGACT 399
Db 181 ACAGCATTTACCTACCGGTACCTCGAAGTGTCCACATCCAGAGGGTGCCTGCCTGACT 240
QY 400 GGCACAACTGCATCAACAATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTGAC 459
Db 241 GGCACAACTGCATCAACAATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTGAC 300
QY 460 TATGTGTTACCGCCCAATCTTCAATGTCTCAGAGTGGCCATCTTTGTAGCAACTGTGCG 519
Db 301 TATGTGTTACCGCCCAATCTTCAATGTCTCAGAGTGGCCATCTTTGTAGCAACTGTGCG 360
QY 520 CCAAAGCTCACATGTTGTCCTCAACTTGGCGGGCCCTTTGGGATCCATTTCGCAACTTGCT 579
Db 361 CCAAAGCTCACATGTTGTCCTCAACTTGGCGGGCCCTTTGGGATCCATTTCGCAACTTGCT 420
QY 580 ATGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTATAA 618
Db 421 ATGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTATAA 459

RESULT 42
BG537816
LOCUS 602566171F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4690768 5',
DEFINITION mRNA sequence.
ACCESSION BG537816
VERSION BG537816.1 GI:13530048
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 658)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1GM1508 row: n column: 17
High quality sequence stop: 643.
Location/Qualifiers
FEATURES
source
1..658
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4690768"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_77"
/notes="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccattggcc); Site 2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor sequence:
5'-ATTCATAGGCGGCGGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
ORIGIN
Query Match 35.9%; Score 457; DB 4; Length 658;
Best Local Similarity 100.0%; Pred. No. 8.2e-234; Indels 0; Gaps 0;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```
QY 812 TTTTCTTCTGCTACAGACATTAATCTTCTGGTCTGTGACTGGTGCATGATGCAGTCTCT 871
Db 457 TTTTCTTCTGCTACAGACATTAATCTTCTGGTCTGTGACTGGTGCATGATGCAGTCTCT 398
QY 872 GTTTTGGCTTTCTACTTTCATCTTGTAGTCTTTAGAGAAAACAGAGAAAATACGATGGTCAACCAGC 931
Db 397 GTTTTGGCTTTCTACTTTCATCTTGTAGTCTTTAGAGAAAACAGAGAAAATACGATGGTCAACCAGC 338
QY 932 AGTTCTTCCGAAATCGTACAGCTGATAGGAAACAGCGAAGCAAGCTGAAATTTTGGTCTTACC 991
Db 337 AGTTCTTCCGAAATCGTACAGCTGATAGGAAACAGCGAAGCAAGCTGAAATTTTGGTCTTACC 278
QY 992 GACTTCAGCTTAAATGGTCTAGGCGGATTTAGTCTTGGGAAGCGACTCTCGATCTATTTC 1051
Db 277 GACTTCAGCTTAAATGGTCTAGGCGGATTTAGTCTTGGGAAGCGACTCTCGATCTATTTC 218
QY 1052 ATGAAGGAATTCGCAACAGCCATTTATGAATAGCGACTCTTAGTCTTTTGACACAGCATTTG 1111
Db 217 ATGAAGGAATTCGCAACAGCCATTTATGAATAGCGACTCTTAGTCTTTTGACACAGCATTTG 158
QY 1112 CACAGCTTTTTCAGAAAATGGCAATTTAGGCAATCAATGTAATTTTCAATGTGTGAA 1171
Db 157 CACAGCTTTTTCAGAAAATGGCAATTTAGGCAATCAATGTAATTTTCAATGTGTGAA 98
QY 1172 ATGGCAATCAACATTTCTGGCCAGCTTTTAAATCTTCAATTTTCAGATTTTCAGAAAATAGGCA 1231
Db 97 ATGGCAATCAACATTTCTGGCCAGCTTTTAAATCTTCAATTTTCAGATTTTCAGAAAATAGGCA 38
QY 1232 CCCATCTGTCTGCCAACCTTAAACCTCTTTCGGTAGGT 1268
Db 37 CCCATCTGTCTGCCAACCTTAAACCTCTTTCGGTAGGT 1
```

```
RESULT 43
LOCUS A1698102 608 bp mRNA linear EST 03-JUN-1999
DEFINITION we20h10.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:2341699 3.
similar to TR:043269 O43269 HSA1H1.1; mRNA sequence.
```

```
A1698102
A1698102
A1698102.1 GI:4986002
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 445.
Location/Qualifiers
1. .608
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2341699"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu24"
/note="Organ: lung; Vector: pRT73D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
```

FEATURES  
source

## ORIGIN

```
Query Match 35.8%; Score 456; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 2.8e-233;
Matches 456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 527 TCACATGTTGTCAAATTGCGGGGCCCTTTGGGATCCATTTGCAACTTGGCTATGGAGA 586
Db 7 TCACATGTTGTCAAATTGCGGGGCCCTTTGGGATCCATTTGCAACTTGGCTATGGAGA 66
QY 587 AAGTGGCTAATTCAGTACTTTTCCCTGTAATATGCTCTTCTGATGTGAATAAATC 646
Db 67 AAGTGGCTAATTCAGTACTTTTCCCTGTAATATGCTCTTCTGATGTGAATAAATC 126
QY 647 TGCCACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCCTTATTCCTGTC 706
Db 127 TGCCACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCCTTATTCCTGTC 186
QY 707 CFTGCCCTGGTCTTCTCTTAATGSCAAGGCTCTCTGGATGCTGTAATGCCCATCTGA 766
Db 187 CFTGCCCTGGTCTTCTCTTAATGSCAAGGCTCTCTGGATGCTGTAATGCCCATCTGA 246
QY 767 TGCATCAGCAATAGTCCATTACACCCCTACAGGAGAGATATAGTTTTCTTGTCTACAG 826
Db 247 TGCATCAGCAATAGTCCATTACACCCCTACAGGAGAGATATAGTTTTCTTGTCTACAG 306
QY 827 ACATTAATCTTCTGCTGCTGTGACTGGTGCATGATGCAAGCTCTCTTTTGGCTTTCACT 886
Db 307 ACATTAATCTTCTGCTGCTGTGACTGGTGCATGATGCAAGCTCTCTTTTGGCTTTCACT 366
QY 887 TCATGTTAGTCTTTAGAGAAACAGGAAAATACGATGGTCCACGACAGTTCTTTCGCAATCG 946
Db 367 TCATGTTAGTCTTTAGAGAAACAGGAAAATACGATGGTCCACGACAGTTCTTTCGCAATCG 426
QY 947 TACAGCTGATAGAACACGCAAGCAAGCTGAAAATT 982
Db 427 TACAGCTGATAGAACACGCAAGCAAGCTGAAAATT 462
```

RESULT 44  
BF839462

```
LOCUS BF839462 487 bp mRNA linear EST 13-JAN-2001
DEFINITION RC4-HT0132-281100-011-h11 HT0132 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF839462
VERSION BF839462.1 GI:12191845
KEYWORDS EST..
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 487)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
```







primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN	Query Match	33.7%;	Score 429;	DB 5;	Length 882;
	Best Local Similarity	100.0%;	Pred. No. 9.8e-219;		
	Matches 429;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	318	AGAAATGAGCGGTGAGACTGCTACAGCATTTACCTACCGGTACCTCGAGTGTCCACCATC	377		
DB	454	AGAAATGAGCGGTGAGACTGCTACAGCATTTACCTACCGGTACCTCGAGTGTCCACCATC	513		
QY	378	CCAGAGGGTGCTGCCCTGACTGCGCAACAACCTGCATCCAAACATGACTTGGCGAGTCTTTTT	437		
DB	514	CCAGAGGGTGCTGCCCTGACTGCGCAACAACCTGCATCCAAACATGACTTGGCGAGTCTTTTT	573		
QY	438	TGAGTGTCCAGTCTGTCTTTGACTATGTGTTACCGCCCATTTCTCAATGTTCAGAGTGGCCA	497		
DB	574	TGAGTGTCCAGTCTGTCTTTGACTATGTGTTACCGCCCATTTCTCAATGTTCAGAGTGGCCA	633		
QY	498	TCTTGTGTTGAGCAACTGTGTGCCCAAGCTCACATGTTGTCCAACTTGC CGGGGCCCTTTT	557		
DB	634	TCTTGTGTTGAGCAACTGTGTGCCCAAGCTCACATGTTGTCCAACTTGC CGGGGCCCTTTT	693		
QY	558	GGGATCCATTGCGAACTTGGCTATGGGAAGTGGCTAAATTCAGTACTTTTTCCCTCTGTAA	617		
DB	694	GGGATCCATTGCGAACTTGGCTATGGGAAGTGGCTAAATTCAGTACTTTTTCCCTCTGTAA	753		
QY	618	ATATGCGTCTTCTGGATGTGAAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAGA	677		
DB	754	ATATGCGTCTTCTGGATGTGAAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAGA	813		
QY	678	GCTCTGTGAGTTTAGGCCCTTAATTCCTGTCCGTGCCCTGGTGCTTCCCTGTAATGGCAAGG	737		
DB	814	GCTCTGTGAGTTTAGGCCCTTAATTCCTGTCCGTGCCCTGGTGCTTCCCTGTAATGGCAAGG	873		
QY	738	CTCTCTCGGA	746		
DB	874	CTCTCTCGGA	882		

```

BP332871      578 bp      mRNA      linear      EST 17-SEP-2004
LOCUS      BP332871 Sugano cDNA library, salivary gland Homo sapiens cDNA
DEFINITION      Clone SLV00570, mRNA sequence.
ACCESSION      BP332871
VERSION      BP332871.1 GI:52262328
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 578)
AUTHORS      Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
      Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE      Sequence comparison of human and mouse genes reveals a homologous
      block structure in the promoter regions
JOURNAL      Genome Res. 14 (9), 1711-1718 (2004)
CONTACT      Yutaka Suzuki
DEPARTMENT      Department of Virology
INSTITUTE      Institute of Medical Science, University of Tokyo
ADDRESS      4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
EMAIL      yezuki@ims.u-tokyo.ac.jp.
FEATURES      Location/Qualifiers
      1..578
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="SLV00570"
      /tissue_type="salivary gland"
      /clone_lib="Sugano cDNA library, salivary gland"

```

ORIGIN	Query Match	33.6%;	Score 428;	DB 5;	Length 578;	
	Best Local Similarity	100.0%;	Pred. No. 3.3e-218;			
	Matches 428;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	835	CTTCTGGTCTGTTGACTGGGTGATGATCAGTCAGTCCTGTGTTTGGCTTTCACCTTCATCTTA	894			
DB	1	CTTCTGGTCTGTTGACTGGGTGATGATCAGTCAGTCCTGTGTTTGGCTTTCACCTTCATCTTA	60			
QY	895	GTCTTAGAGAAAACAGGAAAAATACGATGGTCCACAGCAGTCTTTCGCAATCGTACAGCTG	954			
DB	61	GTCTTAGAGAAAACAGGAAAAATACGATGGTCCACAGCAGTCTTTCGCAATCGTACAGCTG	120			
QY	955	ATAGGACACGCAAGCAAGCTGAAAAATTTTGGCTTACCGACTTCAGCTTAAATGGTCATAGG	1014			
DB	121	ATAGGACACGCAAGCAAGCTGAAAAATTTTGGCTTACCGACTTCAGCTTAAATGGTCATAGG	180			
QY	1015	CGACGATTTGACTTGGGAAGGACCTCTTCGATCTATTTCATGAAGAAATTCGCAAGCCATT	1074			
DB	181	CGACGATTTGACTTGGGAAGGACCTCTTCGATCTATTTCATGAAGAAATTCGCAAGCCATT	240			
QY	1075	ATGAATAGCAGCTGTCTAGTCTTTTGACACAGCAGTTCACAGCTTTTTCGAGAAAAATGGC	1134			
DB	241	ATGAATAGCAGCTGTCTAGTCTTTTGACACAGCAGTTCACAGCTTTTTCGAGAAAAATGGC	300			
QY	1135	AATTTAGGCATCAATGTAACTATTTCATCTGTGTGAAATGGCAATCAAAATTTTCTGGC	1194			
DB	301	AATTTAGGCATCAATGTAACTATTTCATCTGTGTGAAATGGCAATCAAAATTTTCTGGC	360			
QY	1195	CAGTGTTTAAACCTTCAGTTTCACAGAAAAATAGGCACCCATCTGTCTGCCAACCTAAAA	1254			
DB	361	CAGTGTTTAAACCTTCAGTTTCACAGAAAAATAGGCACCCATCTGTCTGCCAACCTAAAA	420			
QY	1255	CTCTTTTCG 1262				
DB	421	CTCTTTTCG 428				
RESULT 49						
LOCUS	BG436513					
DEFINITION	602509189F1 NIH_MGC_79 Homo sapiens CDNA clone IMAGE:4619446 5',	853 bp	mRNA	linear	EST 14-MAR-2001	
ACCESSION	BG436513					
VERSION	BG436513.1	GI:13343019				
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
TITLE	NIH-MGC http://mgc.nci.nih.gov/.					
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)					
COMMENT	Unpublished (1999)					
	Contact: Robert Strausberg, Ph.D.					
	Email: cgapbs-remail.nih.gov					
	Tissue Procurement: CLONTECH Laboratories, Inc.					
	cDNA Library Preparation: CLONTECH Laboratories, Inc.					
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
	DNA Sequencing by: Incyte Genomics, Inc.					
	Clone distribution: MGC clone distribution information can be					
	found through the I.M.A.G.E. Consortium/LLNL at:					
	http://image.llnl.gov					
	Plate: LLCMI379 row: b column: 23					
	High quality sequence stop: 644.					
FEATURES	Location/Qualifiers					
	1..853					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/clone="IMAGE:4619446"					
	/lab host="DH10B (T1 phage-resistant)"					

```

BG436513      853 bp      mRNA      linear      EST 14-MAR-2001
LOCUS      602509189f1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4619446 5',
DEFINITION      mRNA sequence.
ACCESSION      BG436513
VERSION      BG436513.1 GI:13343019
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
1 (bases 1 to 853)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LILCM1379 row: b column: 23
High quality sequence stop: 644.
Location/Qualifiers
1. 853
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4619446"
/lab host="DH10B (T1 phage-resistant)"
FEATURES      source

```

```

/clone_lib="NIH_MGC_79"
Site1=Organ; gccagctctggcc; Vector: pDNR-LIB (Clontech);
Site1: Sfil (ggcagctctggcc); Site2: Sfil
(ggcagctctggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3',
and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.3
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."

```

## ORIGIN

Query Match	32.9%;	Score 419;	DB 4;	Length 853;
Best Local Similarity	100.0%;	Pred. No. 2.3e-213;		
Matches 419;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	558	GGGATCCATTTCGCAATCTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTCGTAA	617	
Db	8	GGGATCCATTTCGCAACTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTCGTAA	67	
QY	618	ATATGCGTCTTCTGGATGTGGAATAAATCTCTGCCACACACAGAAAGACAGACCAATGAAGA	677	
Db	68	ATATGCGTCTTCTGGATGTGGAATAAATCTCTGCCACACACAGAAAGACAGACCAATGAAGA	127	
QY	678	GCNCTGTGAGTTTAGGCCCTTATTCTCTCGGTGCCCTGGTCTTCCTGTAATATGGCAAGG	737	
Db	128	GCNCTGTGAGTTTAGGCCCTTATTCTCTCGGTGCCCTGGTCTTCCTGTAATATGGCAAGG	187	
QY	738	CTCTCTGGATGCTGTAAATGCCCATCTGTATGTCATCAGCATAAAGTCCAATTACAACCCATACA	797	
Db	188	CTCTCTGGATGCTGTAAATGCCCATCTGTATGTCATCAGCATAAAGTCCAATTACAACCCATACA	247	
QY	798	GGAGAGGATATAGTATTTTCTTGCTACAGACATTAATCTTCTCGTGGTCTGTTGACTGGGT	857	
Db	248	GGAGAGGATATAGTATTTTCTTGCTACAGACATTAATCTTCTCGTGGTCTGTTGACTGGGT	307	
QY	858	GATGATGCAGTCTGTTTTCGGCTTTCATCTTCATGTTAGTCTTAGAGAAAAACAGGAAAAATA	917	
Db	308	GATGATGCAGTCTGTTTTCGGCTTTCATCTTCATGTTAGTCTTAGAGAAAAACAGGAAAAATA	367	
QY	918	CGATGGTCAACAGAGTTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTG	976	
Db	368	CGATGGTCAACAGAGTTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTG	426	

**RESULT 50**

BO431458	913 bp	mRNA	linear	EST 24-MAY-2002
LOCUS	AGENCOURT_7836118	NIH_MGC_67	Homo sapiens	cdna clone IMAGE:6153731
DEFINITION	5', mRNA sequence.			
ACCESSION	BO431458			
VERSION	BO431458.1	GI:21170534		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1 (bases 1 to 913)			
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: <a href="mailto:cgabs-r@mail.nih.gov">cgabs-r@mail.nih.gov</a>			
	Tissue procurement: ATCC			
	CDNA Library Preparation: Life Technologies, Inc.			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA sequencing by: Agencourt Bioscience Corporation			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>			
	plate: LLMI13493	row: k	column: 12	

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 00:35:47 ; Search time 246 Seconds  
(without alignments)  
8474.051 Million cell updates/sec

Title: US-10-679-246-1  
Perfect score: 1274  
Sequence: 1 tttcttagtggtttatggt.....ctcttcggtggtggaagc 1274

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 81813859 residues

Word size : 15

Total number of hits satisfying chosen parameters: 4037

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1274	100.0	1274	4	US-09-591-694-1
2	1041	81.7	2440	4	US-09-023-655-588
3	780	61.2	1884	4	US-09-544-618-11
4	21	1.6	139552	4	US-09-949-016-15300
5	20	1.6	28	4	US-09-591-694-42
6	20	1.6	257	4	US-09-016-434-397
7	20	1.6	8342	3	US-08-545-860D-63
8	20	1.6	8342	5	PCT-US94-04486-63
9	20	1.6	8392	1	US-08-080-253-6
10	20	1.6	8392	3	US-08-465-713-6
11	20	1.6	8392	5	PCT-US93-05857-6
12	19	1.5	19	4	US-09-591-694-43
13	19	1.5	345	4	US-09-252-991A-10103
14	19	1.5	601	4	US-09-949-016-19538
15	19	1.5	601	4	US-09-949-016-19539
16	19	1.5	601	4	US-09-949-016-94799
17	19	1.5	601	4	US-09-949-016-94800
18	19	1.5	601	4	US-09-949-016-164402
19	19	1.5	601	4	US-09-949-016-164403
20	19	1.5	636	4	US-09-252-991A-10002
21	19	1.5	1473	4	US-08-277-031B-14
22	19	1.5	1920	4	US-09-252-991A-9817
23	19	1.5	1995	4	US-09-023-655-1055
24	19	1.5	1995	4	US-09-949-016-4610
25	19	1.5	2009	1	US-08-201-118-6
26	19	1.5	2009	2	US-08-238-821B-6
27	19	1.5	2009	5	PCT-US95-05744-6
28	19	1.5	2258	1	US-08-201-118-12
29	19	1.5	2258	2	US-08-238-821B-12
30	19	1.5	2258	4	US-09-023-655-995
31	19	1.5	2258	5	PCT-US95-05744-12
32	19	1.5	3001	4	US-09-539-333D-220
33	19	1.5	4515	4	US-09-252-991A-10037
34	19	1.5	44248	4	US-09-949-016-11829
35	19	1.5	44249	4	US-09-949-016-14885
36	19	1.5	44249	4	US-09-949-016-14491
37	19	1.5	56147	4	US-09-949-016-16352
38	19	1.5	1830121	4	US-09-557-884-1
39	19	1.5	1830121	4	US-09-643-990A-1
40	18	1.4	303	4	US-09-248-796A-7164
41	18	1.4	414	4	US-09-248-796A-852
42	18	1.4	420	4	US-09-621-976-14723
43	18	1.4	451	4	US-09-513-999C-31455
44	18	1.4	493	4	US-09-621-976-11155
45	18	1.4	495	2	US-08-955-848A-46
46	18	1.4	495	5	PCT-US95-03866-1
47	18	1.4	498	5	PCT-US95-03866-17
48	18	1.4	498	5	PCT-US95-03866-35
49	18	1.4	501	5	PCT-US95-03866-19
50	18	1.4	506	2	US-08-955-848A-58
51	18	1.4	506	2	US-08-955-848A-59
52	18	1.4	506	2	US-08-955-848A-60
53	18	1.4	506	2	US-08-955-848A-61
54	18	1.4	506	2	US-08-955-848A-62
55	18	1.4	506	2	US-08-955-848A-63
56	18	1.4	506	2	US-08-955-848A-64
57	18	1.4	506	2	US-08-955-848A-65
58	18	1.4	506	2	US-08-955-848A-66
59	18	1.4	506	2	US-08-955-848A-67
60	18	1.4	506	2	US-08-955-848A-68
61	18	1.4	506	2	US-08-955-848A-69
62	18	1.4	506	2	US-08-955-848A-70
63	18	1.4	506	2	US-08-955-848A-71
64	18	1.4	506	2	US-08-955-848A-72
65	18	1.4	506	2	US-08-955-848A-73
66	18	1.4	506	2	US-08-955-848A-74
67	18	1.4	506	2	US-08-955-848A-75
68	18	1.4	506	2	US-08-955-848A-76
69	18	1.4	506	2	US-08-955-848A-77
70	18	1.4	506	2	US-08-955-848A-78
71	18	1.4	506	2	US-08-955-848A-79
72	18	1.4	506	2	US-08-955-848A-80
73	18	1.4	506	2	US-08-955-848A-81
74	18	1.4	507	2	US-08-955-848A-82
75	18	1.4	507	2	US-08-955-848A-83
76	18	1.4	507	2	US-08-955-848A-84
77	18	1.4	507	2	US-08-955-848A-85
78	18	1.4	507	2	US-08-955-848A-86
79	18	1.4	507	2	US-08-955-848A-87
80	18	1.4	507	2	US-08-955-848A-88
81	18	1.4	507	2	US-08-955-848A-89
82	18	1.4	507	2	US-08-955-848A-90
83	18	1.4	512	2	US-08-628-428-3
84	18	1.4	512	2	US-08-628-428-6
85	18	1.4	555	4	US-09-702-705-380
86	18	1.4	555	4	US-09-736-457-380
87	18	1.4	555	4	US-09-614-124B-380
88	18	1.4	555	4	US-09-671-325-380
89	18	1.4	555	4	US-09-589-184-380
90	18	1.4	555	4	US-09-658-824-380
91	18	1.4	601	4	US-09-949-016-86177
92	18	1.4	601	4	US-09-949-016-119438
93	18	1.4	601	4	US-09-949-016-145245
94	18	1.4	601	4	US-09-949-016-188597
95	18	1.4	601	4	US-09-949-016-188598
96	18	1.4	601	4	US-09-949-016-205578
97	18	1.4	820	3	US-08-836-252A-5
98	18	1.4	820	3	US-08-482-918-45
99	18	1.4	820	3	US-09-224-681-45
100	18	1.4	820	3	US-08-336-728A-45

Sequence 12, Appl  
Sequence 12, Appl  
Sequence 995, Appl  
Sequence 12, Appl  
Sequence 220, Appl  
Sequence 10037, A  
Sequence 11829, A  
Sequence 14485, A  
Sequence 14491, A  
Sequence 16352, A  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 7164, Appl  
Sequence 852, Appl  
Sequence 14723, A  
Sequence 11455, A  
Sequence 11155, A  
Sequence 46, Appl  
Sequence 1, Appl  
Sequence 17, Appl  
Sequence 35, Appl  
Sequence 19, Appl  
Sequence 58, Appl  
Sequence 59, Appl  
Sequence 60, Appl  
Sequence 61, Appl  
Sequence 62, Appl  
Sequence 63, Appl  
Sequence 64, Appl  
Sequence 65, Appl  
Sequence 66, Appl  
Sequence 67, Appl  
Sequence 68, Appl  
Sequence 69, Appl  
Sequence 70, Appl  
Sequence 71, Appl  
Sequence 72, Appl  
Sequence 73, Appl  
Sequence 74, Appl  
Sequence 75, Appl  
Sequence 76, Appl  
Sequence 77, Appl  
Sequence 78, Appl  
Sequence 79, Appl  
Sequence 80, Appl  
Sequence 81, Appl  
Sequence 47, Appl  
Sequence 48, Appl  
Sequence 49, Appl  
Sequence 50, Appl  
Sequence 51, Appl  
Sequence 52, Appl  
Sequence 53, Appl  
Sequence 54, Appl  
Sequence 55, Appl  
Sequence 3, Appl  
Sequence 6, Appl  
Sequence 380, Appl  
Sequence 380, Appl  
Sequence 380, Appl  
Sequence 380, Appl  
Sequence 380, Appl  
Sequence 86177, A  
Sequence 139438, A  
Sequence 145245, A  
Sequence 188597, A  
Sequence 188598, A  
Sequence 205578, A  
Sequence 5, Appl  
Sequence 45, Appl  
Sequence 45, Appl  
Sequence 45, Appl















Sequence 14590, A  
Sequence 56, Appl  
Sequence 56, Appl  
Sequence 3, Appl  
Sequence 15716, A  
Sequence 12451, A  
Sequence 17359, A  
Sequence 16961, A  
Sequence 15945, A  
Sequence 16949, A  
Sequence 15301, A  
Sequence 17555, A  
Sequence 12304, A  
Sequence 13787, A  
Sequence 11991, A  
Sequence 17249, A  
Sequence 13832, A  
Sequence 13627, A  
Sequence 16102, A  
Sequence 12633, A  
Sequence 15585, A  
Sequence 12832, A  
Sequence 16704, A  
Sequence 13846, A

16 1.3 29615 4 US-09-949-016-14590  
16 1.3 30246 4 US-08-956-171E-56  
16 1.3 30246 4 US-08-781-986A-56  
16 1.3 30350 4 US-10-118-328-3  
16 1.3 30360 4 US-09-949-016-15716  
16 1.3 30402 4 US-09-949-016-12451  
16 1.3 31407 4 US-09-949-016-17359  
16 1.3 31602 4 US-09-949-016-16961  
16 1.3 31623 4 US-09-949-016-15945  
16 1.3 33272 4 US-09-949-016-16949  
16 1.3 33624 4 US-09-949-016-15301  
16 1.3 34027 4 US-09-949-016-17555  
16 1.3 34628 4 US-09-949-016-12304  
16 1.3 34779 4 US-09-949-016-13787  
16 1.3 35256 4 US-09-949-016-11991  
16 1.3 35337 4 US-09-949-016-17249  
16 1.3 35630 4 US-09-949-016-13832  
16 1.3 36311 4 US-09-949-016-13627  
16 1.3 36661 4 US-09-949-016-16102  
16 1.3 36907 4 US-09-949-016-12633  
16 1.3 36913 4 US-09-949-016-15585  
16 1.3 37711 4 US-09-949-016-12832  
16 1.3 37712 4 US-09-949-016-16704  
16 1.3 37715 4 US-09-949-016-13846

ALIGNMENTS

RESULT 1  
US-09-591-694-1  
; Sequence 1, Application US/09591694  
; Patent No. 6638734  
; GENERAL INFORMATION:  
; APPLICANT: John C. Reed  
; APPLICANT: Shu-ichi Matsuzawa  
; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved  
; FILE OF INVENTION: in Protein Degradation, Products and Methods Related Thereto  
; FILE REFERENCE: P-LJ 4220  
; CURRENT APPLICATION NUMBER: US/09/591,694  
; CURRENT FILING DATE: 2000-06-09  
; EARLIER APPLICATION NUMBER: US 09/330,517  
; EARLIER FILING DATE: 1999-06-11  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1274  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (274)...(1167)  
US-09-591-694-1

Query Match 100.0%; Score 1274; DB 4; Length 1274;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTTCTTTAGTTGTTATGGTCCATTTTCTATTTTACGATTTATTTCTATGTTAGTCTAT 60  
1 TTTCTTTAGTTGTTATGGTCCATTTTCTATTTTACGATTTATTTCTATGTTAGTCTAT 60

61 CCAAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 120  
61 CCAAAGACGATTAAAGGAGTTCCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 120

121 CAGTGTACAGATCCTAATAAAGTGCAATTCAGTGTAAATTTTATTTTAAATATCTTTT 180  
121 CAGTGTACAGATCCTAATAAAGTGCAATTCAGTGTAAATTTTATTTTAAATATCTTTT 180

181 TTAATCCTATTTTCTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
181 TTAATCCTATTTTCTCTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240

241 TATGCGCATGTAACCATTTATTTATAAAGTAAAGTCAATGTTTATTAATTTTCTCTGCT 300  
241 TATGCGCATGTAACCATTTATTTATAAAGTAAAGTCAATGTTTATTAATTTTCTCTGCT 300

301 CCTTATGTTATTTATTTTATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
301 CCTTATGTTATTTATTTTATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360

361 TCGAAGTGTCCACCATCCAGAGGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
361 TCGAAGTGTCCACCATCCAGAGGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

421 GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTTGTGATGTTTGAATGTTTACCGCCCA 480  
421 GACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGCTTGTGATGTTTGAATGTTTACCGCCCA 480

481 CAATGTGAGAGTGGCCATCTTTGTTGTTAGCAACTGTGCTGCTGCTGCTGCTGCTGCTGCT 540  
481 CAATGTGAGAGTGGCCATCTTTGTTGTTAGCAACTGTGCTGCTGCTGCTGCTGCTGCTGCT 540

541 ACTTGGCGGGGCTTTTGGAGTGTCCAGTCTGCTTGTGATGTTTGAATGTTTACCGCCCA 600  
541 ACTTGGCGGGGCTTTTGGAGTGTCCAGTCTGCTTGTGATGTTTGAATGTTTACCGCCCA 600

601 GTACTTTTCCCTGTAAATATGCGTCTTCTGATGTTGAAATTAATCTGCTGCTGCTGCTGCT 660  
601 GTACTTTTCCCTGTAAATATGCGTCTTCTGATGTTGAAATTAATCTGCTGCTGCTGCTGCT 660

661 AAAGCAGACCATTAAGAGCTCTGCTGAGTGTAGGCTTATTTCTGCTGCTGCTGCTGCTGCT 720  
661 AAAGCAGACCATTAAGAGCTCTGCTGAGTGTAGGCTTATTTCTGCTGCTGCTGCTGCTGCT 720

721 TCCTGTAATGCGCAAGGCTCTGCTGATGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTG 780  
721 TCCTGTAATGCGCAAGGCTCTCTGCTGATGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCT 780

781 TCCATTACCAACCTACAGGGAGAGATATAGTTTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
781 TCCATTACCAACCTACAGGGAGAGATATAGTTTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 840

841 GGTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
841 GGTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

901 GAGAAACAGGAAATATACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
901 GAGAAACAGGAAATATACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

961 ACACGCAAGAGCTGAAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
961 ACACGCAAGAGCTGAAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

1021 TTGACTTGGGAAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1021 TTGACTTGGGAAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

1081 AGCGAGTGTCTAGTCTTTTGACACCCAGCATTTGACAGCTTTTTCGCAAAATGGCAATTTA 1140  
1081 AGCGAGTGTCTAGTCTTTTGACACCCAGCATTTGACAGCTTTTTCGCAAAATGGCAATTTA 1140

1141 GGCATCAATGTAATTTTCCATGTTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
1141 GGCATCAATGTAATTTTCCATGTTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200

1201 TTAATACTTCAGTTTTCACAGAAATTAAGGCAACCCATCTGCTGCTGCTGCTGCTGCTGCT 1260  
1201 TTAATACTTCAGTTTTCACAGAAATTAAGGCAACCCATCTGCTGCTGCTGCTGCTGCTGCT 1260

1261 CGGTAGGTGGAAGC 1274  
1261 CGGTAGGTGGAAGC 1274



RESULT 2  
US-09-023-655-588  
; Sequence 588, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 588:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2440 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: SPLNFT01  
; CLONE: 29244  
; US-09-023-655-588

Query Match 81.7%; Score 1041; DB 4; Length 2440;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 234 AAGGACTTATGGCATGTAACCATTTATTAAGTAAGTCATGGTTATTAATTATTTTCT 293  
DB 467 AAGGACTTATGGCATGTAACCATTTATTAAGTAAGTCATGGTTATTAATTATTTTCT 526  
QY 294 CCGTCCCTCTTATGTATTTATTTTACAAATGAGCCGTGACAGCTGCTACAGCACTTACCTAC 353  
DB 527 CCGTCCCTCTTATGTATTTATTTTACAAATGAGCCGTGACAGCTGCTACAGCACTTACCTAC 586  
QY 354 CGGTACTCTGAAAGTGTCCACCATCCAGAGGGTGGCTGCTGACCTGACCTGACCACTGCATC 413  
DB 587 CGGTACTCTGAAAGTGTCCACCATCCAGAGGGTGGCTGCTGACCTGACCTGACCACTGCATC 646  
QY 414 CAACAAATGACTGGCGAGTCTTTTTCAGTGTCCAGTGTCCAGTGTCTGTTGACTATGTGTTACCGCC 473  
DB 647 CAACAAATGACTGGCGAGTCTTTTTCAGTGTCCAGTGTCTGTTGACTATGTGTTACCGCC 706  
QY 474 CATTCTTCAATGTACAGTGGCCATCTTTTGTAGCAACTGTGCGCCCAAGCTCACATG 533  
DB 707 CATTCTTCAATGTACAGTGGCCATCTTTTGTAGCAACTGTGCGCCCAAGCTCACATG 766

QY 534 TTGTCCAACTTGCCTGGGGCCCTTTTGGGATCCATTTCGCAACTTGGCTATGAGAAAAGTGGC 593  
DB 767 TTGTCCAACTTGCCTGGGGCCCTTTTGGGATCCATTTCGCAACTTGGCTATGAGAAAAGTGGC 826  
QY 594 TAAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATTAACCTGTCACCA 653  
DB 827 TAAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAATTAACCTGTCACCA 886  
QY 654 CACAGAAAAGCAGACCAATGAAGAGCTCTGTGAGTTTAGGCTTATTCCTGTCCGTGCCC 713  
DB 887 CACAGAAAAGCAGACCAATGAAGAGCTCTGTGAGTTTAGGCTTATTCCTGTCCGTGCCC 946  
QY 714 TGGTGTCTTCTGTAAATGCAAGGCTCTCTGATGCTGAATGCCCCCATCTGATGATCA 773  
DB 947 TGGTGTCTTCTGTAAATGCAAGGCTCTCTGATGCTGAATGCCCCCATCTGATGATCA 1006  
QY 774 GCATAGTCCATTACAAACCTTACAGGAGAGGATATAGTTTTCTTGTACACACATTAA 833  
DB 1007 GCATAGTCCATTACAAACCTTACAGGAGAGGATATAGTTTTCTTGTACACACATTAA 1066  
QY 834 TCTTCTGTGTCTTGTGACTGGGTGATGATGAGTCTCTGTCTTGTGGCTTTTCACTTTCATGTT 893  
DB 1067 TCTTCTGTGTCTTGTGACTGGGTGATGATGAGTCTCTGTCTTGTGGCTTTTCACTTTCATGTT 1126  
QY 894 AGTCTTAGAGAAAACAGGAAAAATACGATGCTCACAGCAGTCTTTCGCAATGTCAGCT 953  
DB 1127 AGTCTTAGAGAAAACAGGAAAAATACGATGCTCACAGCAGTCTTTCGCAATGTCAGCT 1186  
QY 954 GATAGAACCGCAGCAGCTGAAAAATTTTGTACCGACTTACGCTAAATGCTCATAG 1013  
DB 1187 GATAGAACCGCAGCAGCTGAAAAATTTTGTACCGACTTACGCTAAATGCTCATAG 1246  
QY 1014 GCGACGATTGACTTGGGAAGCAGCTCTCGATCTATTATGAGGAATTTGCAACAGCCAT 1073  
DB 1247 GCGACGATTGACTTGGGAAGCAGCTCTCGATCTATTATGAGGAATTTGCAACAGCCAT 1306  
QY 1074 TATGAATAGCGACTGTCTAGTCTTTGACACACAGCATGTCACAGCTTTTTCGAGAAAATGG 1133  
DB 1307 TATGAATAGCGACTGTCTAGTCTTTGACACACAGCATGTCACAGCTTTTTCGAGAAAATGG 1366  
QY 1134 CAATTTAGCATCAATGTAACCTATTTCATGCTGTTGAAATGCAATCAACATTTTCTGG 1193  
DB 1367 CAATTTAGCATCAATGTAACCTATTTCATGCTGTTGAAATGCAATCAACATTTTCTGG 1426  
QY 1194 CCAGTGTTTTAAACTTCAGTCTTTCAGAAAATAAGGCAACCCATCTCTGCGCAACCTAAA 1253  
DB 1427 CCAGTGTTTTAAACTTCAGTCTTTCAGAAAATAAGGCAACCCATCTCTGCGCAACCTAAA 1486  
QY 1254 ACTCTTTCGGTAGTGGGAAGC 1274  
DB 1487 ACTCTTTCGGTAGTGGGAAGC 1507

RESULT 3  
US-09-544-618-11  
; Sequence 11, Application US/09544618  
; Patent No. 6503502  
; GENERAL INFORMATION:  
; APPLICANT: Telemann, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Cohen, Daniel  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, PROTEINS, DRUGS AND DIAGNOSTIC  
; TITLE OF INVENTION: AGENTS OF USE IN TREATING CANCER  
; FILE REFERENCE: 065691-0139  
; CURRENT APPLICATION NUMBER: US/09/544,618  
; CURRENT FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 1884  
; TYPE: DNA  
; ORGANISM: TSAP 3





COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 397:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 257 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRAITUT03  
CLONE: 2113436  
US-09-016-434-397

Query Match 1.6%; Score 20; DB 4; Length 257;  
Best Local Similarity 100.0%; Pred.No. 6.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 850 GACTGGGTGATGATGCAGTC 869  
DB 222 GACTGGGTGATGATGCAGTC 241

RESULT 7  
US-08-545-860D-63/c  
Sequence 63, Application US/08545860D  
Patent No. 6040140  
GENERAL INFORMATION:  
APPLICANT: Croce, Carlo  
APPLICANT: Canaan, Eli  
TITLE OF INVENTION: Diagnostics, Therapeutics and Methods  
TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias  
TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &  
ADDRESSEE: No. 6040140ris  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/545,860D  
FILING DATE: 07-MAR-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/04496  
FILING DATE: 22-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/10930

FILING DATE: 09-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/327,392  
FILING DATE: 19-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/320,559  
FILING DATE: 11-OCT-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/062,443  
FILING DATE: 14-MAY-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/971,094  
FILING DATE: 30-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/888,839  
FILING DATE: 27-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/805,093  
FILING DATE: 11-DEC-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Deluca Esq., Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: TJU-1262  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8342 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2..265  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 595..666  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 2353..2484  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3032..3145  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 6788..6934  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7967..8062  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 8304..8342  
US-08-545-860D-63

Query Match 1.6%; Score 20; DB 3; Length 8342;  
Best Local Similarity 100.0%; Pred.No. 7.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TGTAAATTTTATTTTAAAT 173  
DB 6400 TGTAAATTTTATTTTAAAT 6381

RESULT 8  
PCT-US94-04496-63/c  
Sequence 63, Application PC/TUS9404496  
GENERAL INFORMATION:  
APPLICANT: Croce, Carlo  
APPLICANT: Canaan, Eli

;; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods  
;; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias  
;; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1  
;; NUMBER OF SEQUENCES: 86  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &  
;; ADDRESS: Norris  
;; STREET: One Liberty Place, 46th floor  
;; CITY: Philadelphia  
;; STATE: Pennsylvania  
;; COUNTRY: USA  
;; ZIP: 19103  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US94/04496  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Deluca Esq., Mark  
;; REGISTRATION NUMBER: 33,229  
;; REFERENCE/DOCKET NUMBER: TJU-1242  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (215) 568-3100  
;; TELEFAX: (215) 568-3439  
;; INFORMATION FOR SEQ ID NO: 63:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 8342 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 2..265  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 595..666  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 2353..2484  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 3032..3145  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 6788..6934  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 7967..8062  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 8304..8342  
;; PCT-US94-04496-63  
  
Query Match 1.6%; Score 20; DB 5; Length 8342;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 154 TGTAAATTTATTTTAAAT 173  
Db 6400 TGTAAATTTATTTTAAAT 6381  
  
RESULT 9  
US-08-080-255-6/c  
; Sequence 6, Application US/08080255  
; Patent No. 5487970

;; GENERAL INFORMATION:  
;; APPLICANT: Rowley, Janet D.  
;; APPLICANT: Diaz, Manuel O.  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
;; TITLE OF INVENTION: DETECTING GENE REARRANGEMENTS AND  
;; TITLE OF INVENTION: TRANSLOCATIONS  
;; NUMBER OF SEQUENCES: 8  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Arnold, White & Durkee  
;; STREET: P. O. Box 4433  
;; CITY: Houston  
;; STATE: Texas  
;; COUNTRY: USA  
;; ZIP: 77210  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/080,255  
;; FILING DATE: 19930617  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Parker, David L.  
;; REGISTRATION NUMBER: 32,165  
;; REFERENCE/DOCKET NUMBER: ARCD:072/PAR  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (512) 320-7200  
;; TELEFAX: (512) 474-7577  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 8392 base pairs  
;; TYPE: NUCLEIC ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; US-08-080-255-6  
  
Query Match 1.6%; Score 20; DB 1; Length 8392;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 154 TGTAAATTTATTTTAAAT 173  
Db 6450 TGTAAATTTATTTTAAAT 6431  
  
RESULT 10  
US-08-465-713-6/c  
; Sequence 6, Application US/08465713  
; Patent No. 6121419  
;; GENERAL INFORMATION:  
;; APPLICANT: Rowley, Janet D.  
;; APPLICANT: Diaz, Manuel O.  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
;; TITLE OF INVENTION: DETECTING GENE REARRANGEMENTS AND  
;; TITLE OF INVENTION: TRANSLOCATIONS  
;; NUMBER OF SEQUENCES: 8  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Arnold, White & Durkee  
;; STREET: P. O. Box 4433  
;; CITY: Houston  
;; STATE: Texas  
;; COUNTRY: USA  
;; ZIP: 77210  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/465,713

;; FILING DATE: 06-JUN-1995  
;; CLASSIFICATION: 530  
;; PRIOR APPLICATION NUMBER: US/08/080,255  
;; FILING DATE: 17 JUNE 1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Parker, David L.  
;; REGISTRATION NUMBER: 32,165  
;; REFERENCE/DOCKET NUMBER: ARCD:072/PAR  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (512) 320-7200  
;; TELEFAX: (512) 474-7577  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 8392 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
US-08-465-713-6

Query Match 1.6%; Score 20; DB 3; Length 8392;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TGTAAATTTTATTTTAAAT 173  
DB 6450 TGTAAATTTTATTTTAAAT 6431

RESULT 11  
PCT-US93-05857-6/c  
;; Sequence 6, Application PC/TUS9305857  
;; GENERAL INFORMATION:  
;; APPLICANT: Board of Regents  
;; APPLICANT: The University of Texas System  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING  
;; TITLE OF INVENTION: GENE REARRANGEMENTS AND TRANSLOCATIONS  
;; NUMBER OF SEQUENCES: 8  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Arnold, White & Durkee  
;; STREET: P. O. Box 4433  
;; CITY: Houston  
;; STATE: Texas  
;; COUNTRY: USA  
;; ZIP: 77210  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US93/05857  
;; FILING DATE: 19930617  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/900,689  
;; FILING DATE: 17/06/92  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Parker, David L.  
;; REGISTRATION NUMBER: 32,165  
;; REFERENCE/DOCKET NUMBER: ARCD:072/PAR  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (512) 320-7200  
;; TELEFAX: (512) 474-7577  
;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 8392 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
PCT-US93-05857-6

Query Match 1.6%; Score 20; DB 5; Length 8392;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TGTAAATTTTATTTTAAAT 173  
DB 6450 TGTAAATTTTATTTTAAAT 6431

RESULT 12  
US-09-591-694-43/c  
;; Sequence 43, Application US/09591694  
;; Patent No. 6638734  
;; GENERAL INFORMATION:  
;; APPLICANT: Shu-Tchi Matsuzawa  
;; APPLICANT: John C. Reed  
;; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved  
;; TITLE OF INVENTION: in Protein Degradation, Products and Methods Related Thereto  
;; FILE REFERENCE: P-LJ 4220  
;; CURRENT APPLICATION NUMBER: US/09/591,694  
;; EARLIER FILING DATE: 2000-06-09  
;; EARLIER APPLICATION NUMBER: US 09/330,517  
;; EARLIER FILING DATE: 1999-06-11  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 43  
;; LENGTH: 19  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Primer  
US-09-591-694-43

Query Match 1.5%; Score 19; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 562 TCCATTGCGCAACTTGGCTA 580  
DB 19 TCCATTGCGCAACTTGGCTA 1

RESULT 13  
US-09-252-991A-10103  
;; Sequence 10103, Application US/09252991A  
;; Patent No. 6551795  
;; GENERAL INFORMATION:  
;; APPLICANT: Marc J. Rubenfield et al.  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
;; FILE REFERENCE: 107196.136  
;; CURRENT APPLICATION NUMBER: US/09/252,991A  
;; CURRENT FILING DATE: 1999-02-18  
;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 10103  
;; LENGTH: 345  
;; TYPE: DNA  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-10103

Query Match 1.5%; Score 19; DB 4; Length 345;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 918 CGATGGTCACGACGATTC 936  
DB 98 CGATGGTCACGACGATTC 116

```
RESULT 14
US-09-949-016-19538
; Sequence 19538, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19538
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-19538

Query Match          1.5%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GTAATTTTATTTTAAAT 173
Db 40 GTAATTTTATTTTAAAT 58

RESULT 15
US-09-949-016-19539
; Sequence 19539, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19539
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-19539

Query Match          1.5%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GTAATTTTATTTTAAAT 173
Db 241 GTAATTTTATTTTAAAT 259

RESULT 16
US-09-949-016-94799
; Sequence 94799, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94799
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-94799

Query Match          1.5%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GTAATTTTATTTTAAAT 173
Db 40 GTAATTTTATTTTAAAT 58

RESULT 17
US-09-949-016-94800
; Sequence 94800, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94800
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-94800

Query Match          1.5%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GTAATTTTATTTTAAAT 173
Db 241 GTAATTTTATTTTAAAT 259

RESULT 18
US-09-949-016-164402
; Sequence 164402, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```



```
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 972 AGCTGAAATTTTGTCTAC 990
Db 699 AGCTGAAATTTTGTCTAC 717
|||||
RESULT 22
US-09-252-991A-9817/c
; Sequence 9817, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 9817
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9817
Query Match 1.5%; Score 19; DB 4; Length 1920;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 918 CGATGTCACCACAGTTC 936
Db 1720 CGATGTCACCACAGTTC 1702
|||||
RESULT 23
US-09-023-655-1055
; Sequence 1055, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1055:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1995 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g181299
US-09-023-655-1055
Query Match 1.5%; Score 19; DB 4; Length 1995;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 972 AGCTGAAATTTTGTCTAC 990
Db 898 AGCTGAAATTTTGTCTAC 916
|||||
RESULT 24
US-09-949-016-4610
; Sequence 4610, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4610
; LENGTH: 1995
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4610
Query Match 1.5%; Score 19; DB 4; Length 1995;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 972 AGCTGAAATTTTGTCTAC 990
Db 898 AGCTGAAATTTTGTCTAC 916
|||||
RESULT 25
US-08-201-118-6
; Sequence 6, Application US/08201118
; Patent No. 5786191
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; TITLE OF INVENTION: CLONING AND EXPRESSION OF COMPLEMENTARY
; DNA FOR MULTIPLE MEMBERS OF THE HUMAN CYTOCHROME P450 2C
; TITLE OF INVENTION: SUBFAMILY
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
```



; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/201,118  
; FILING DATE: 22-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962  
; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Liebeschuetz, Joe  
; REGISTRATION NUMBER: 37,505  
; REFERENCE/DOCKET NUMBER: 15280-192-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2009 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-201-118-6

Query Match 1.5%; Score 19; DB 1; Length 2009;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 972 AGCTGAAATTTTGCTTAC 990  
|||||  
DB 898 AGCTGAAATTTTGCTTAC 916  
|||||

RESULT 26  
US-08-238-821B-6  
; Sequence 6, Application US/08238821B  
; Patent No. 5912120  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; APPLICANT: ROMKES-SPARKS, Marjorie  
; APPLICANT: DE MORAIS, Sonia M.F.  
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-  
; TITLE OF INVENTION: MEPHENYTOIN METABOLISM  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/238,821B  
; FILING DATE: 06-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/201,118  
; FILING DATE: 22-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962

; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Liebeschuetz, Joe  
; REGISTRATION NUMBER: 37,505  
; REFERENCE/DOCKET NUMBER: 15280-192110US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 326-2400  
; TELEFAX: (650) 326-2422  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2009 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..199  
; OTHER INFORMATION: /note= "Corresponds to positions -199to -  
; OTHER INFORMATION: 1 for 29c of Figure 2."  
US-08-238-821B-6

Query Match 1.5%; Score 19; DB 2; Length 2009;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 972 AGCTGAAATTTTGCTTAC 990  
|||||  
DB 898 AGCTGAAATTTTGCTTAC 916  
|||||

RESULT 27  
PCT-US95-05744-6  
; Sequence 6, Application PC/TUS9505744  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; APPLICANT: ROMKES-SPARKS, Marjorie  
; APPLICANT: DE MORAIS, Sonia M.F.  
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT  
; TITLE OF INVENTION: OF S-MEPHNYTOIN METABOLISM  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05744  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/238,821  
; FILING DATE: 06-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/201,118  
; FILING DATE: 22-FEB-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962  
; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 15280-192-1-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400

RESULT 29  
US-08-238-821B-12  
; Sequence 12, Application US/08238821B  
; Patent No. 5912120  
; GENERAL INFORMATION:  
; APPLICANT: GOLDSTEIN, Joyce A.  
; APPLICANT: ROMKES-SPARKS, Marjorie  
; APPLICANT: DE MORAIS, Sonia M.F.  
; TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN  
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-  
; TITLE OF INVENTION: MEPHENYTOIN METABOLISM  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/238,821B  
; FILING DATE: 06-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/201,118  
; FILING DATE: 22-FEB-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,962  
; FILING DATE: 09-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Liebeschuetz, Joe  
; REGISTRATION NUMBER: 37,505  
; REFERENCE/DOCKET NUMBER: 15280-192110US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 326-2400  
; TELEFAX: (650) 326-2422  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2258 base.pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..18  
; OTHER INFORMATION: /note= "Corresponds to positions -18 to-1  
; OTHER INFORMATION: for 6b of Figure 2."  
US-08-238-821B-12

Query Match 1.5%; Score 19; DB 2; Length 2258;  
Best Local Similarity 100.0%; Pred.No.23;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 AGCTGAAATTTTGCTTAC 990  
Db 741 AGCTGAAATTTTGCTTAC 759

RESULT 30  
US-09-023-655-995  
; Sequence 995, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart

```
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 995:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2258 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g171181
;
US-09-023-655-995

Query Match 1.5%; Score 19; DB 4; Length 2258;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 972 AGCTGAAATTTTGCTTAC 990
|||||
Db 741 AGCTGAAATTTTGCTTAC 759

RESULT 31
PCT-US95-05744-12
; Sequence 12, Application PC/TUS9505744
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Joyce A.
; APPLICANT: ROMKES-SPARKS, Marjorie
; APPLICANT: DE MORAIS, Sonia M.F.
; TITLE OF INVENTION: CLONING EXPRESSION AND DIAGNOSIS OF HUMAN
; TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05744
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/238,821
; FILING DATE: 06-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/201,118
; FILING DATE: 22-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,962
; FILING DATE: 09-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 15280-192-1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2258 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
PCT-US95-05744-12

Query Match 1.5%; Score 19; DB 5; Length 2258;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 972 AGCTGAAATTTTGCTTAC 990
|||||
Db 741 AGCTGAAATTTTGCTTAC 759

RESULT 32
US-09-539-333D-220
; Sequence 220, Application US/09539333D
; Patent No. 6476208
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: GENSET.047AUS
; CURRENT APPLICATION NUMBER: US/09/539,333D
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: US 09/416,384
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 231
```

```
; SOFTWARE: Patent.pm
; SEQ ID NO 220
; LENGTH: 3001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1501
; OTHER INFORMATION: 99-27335-191 : polymorphic base A or C
; FEATURE:
; NAME/KEY: misc.binding
; LOCATION: 1481..1500
; OTHER INFORMATION: 99-27335-191.misl,
; FEATURE:
; NAME/KEY: misc.binding
; LOCATION: 1502..1521
; OTHER INFORMATION: 99-27335-191.mis2, complement
; FEATURE:
; NAME/KEY: primer.bnd
; LOCATION: 1322..1342
; OTHER INFORMATION: upstream amplification primer
; FEATURE:
; NAME/KEY: primer.bnd
; LOCATION: 1768..1788
; OTHER INFORMATION: downstream amplification primer, complement
; FEATURE:
; NAME/KEY: misc.binding
; LOCATION: 1489..1513
; OTHER INFORMATION: 99-27335-191 probe
; US-09-539-333D-220

Query Match 1.5%; Score 19; DB 4; Length 3001;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 156 TAATTTTATTTTAAATA 174
Db 765 TAATTTTATTTTAAATA 783

RESULT 33
US-09-252-991A-10037
; Sequence 10037, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10037
; LENGTH: 4515
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-10037

Query Match 1.5%; Score 19; DB 4; Length 4515;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 918 CGATGTCACACAGTTC 936
Db 240 CGATGTCACACAGTTC 258

RESULT 34
US-09-949-016-11829/c
; Sequence 11829, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11829
; LENGTH: 44248
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-11829

Query Match 1.5%; Score 19; DB 4; Length 44248;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GTAATTTTATTTTAAAT 173
Db 9413 GTAATTTTATTTTAAAT 9395

RESULT 35
US-09-949-016-14485/c
; Sequence 14485, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14485
; LENGTH: 44249
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14485

Query Match 1.5%; Score 19; DB 4; Length 44249;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GTAATTTTATTTTAAAT 173
Db 9413 GTAATTTTATTTTAAAT 9395

RESULT 36
US-09-949-016-14491/c
; Sequence 14491, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14491
; LENGTH: 44249
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14491

Query Match
Best Local Similarity 100.0%; Pred. No. 26; Length 44249;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GTAAATTTATTTTAAAT 173
DB 9413 GTAAATTTATTTTAAAT 9395

RESULT 37
US-09-949-016-16352
; Sequence 16352, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16352
; LENGTH: 56147
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16352

Query Match
Best Local Similarity 100.0%; Pred. No. 26; Length 56147;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 972 AGCTGAAAATTTTGCTTAC 990
DB 25216 AGCTGAAAATTTTGCTTAC 25234

RESULT 38
US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
```

```
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match
Best Local Similarity 100.0%; Pred. No. 29; Length 1830121;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 586 AAAGTGGCTAATTCAGTAC 604
DB 1402767 AAAGTGGCTAATTCAGTAC 1402749

RESULT 39
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186PIC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match      1.4%; Score 19; DB 4; Length 1830121;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      586 AAAGTGCTAAATTCAGTAC 604
DB      1402767 AAAGTGCTAAATTCAGTAC 1402749

RESULT 40
US-09-248-796A-7164/c
; Sequence 7164, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7164
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-7164

Query Match      1.4%; Score 18; DB 4; Length 303;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30 ATTTAGCATTTATTATT 47
DB      47 ATTTAGCATTTATTATT 30

RESULT 41
US-09-248-796A-852
; Sequence 852, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
```

```
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 852
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-852

Query Match      1.4%; Score 18; DB 4; Length 414;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      167 TTTTAATATCTTTTTTAA 184
DB      389 TTTTAATATCTTTTTTAA 406

RESULT 42
US-09-621-976-14723
; Sequence 14723, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 14723
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-14723

Query Match      1.4%; Score 18; DB 4; Length 420;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      156 TAAATTTTATTTTAAAT 173
DB      55 TAAATTTTATTTTAAAT 72

RESULT 43
US-09-513-999C-31455
; Sequence 31455, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 31455
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 119
; OTHER INFORMATION: r=a or g
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 376
; OTHER INFORMATION: n=a, g, c o r t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 382
; OTHER INFORMATION: s=g o r c
US-09-513-999C-31455

Query Match      1.4%; Score 18; DB 4; Length 451;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 TTATGATTTTATTTTCAGA 320
DB 222 TTATGATTTTATTTTCAGA 239

RESULT 44
US-09-621-976-11155
; Sequence 11155, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 11155
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-11155

Query Match      1.4%; Score 18; DB 4; Length 493;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 TTTCATTTTAGCATTTA 42
DB 277 TTTCATTTTAGCATTTA 294

RESULT 45
US-09-955-848A-46/c
; Sequence 46, Application US/08955848A
; Patent No. 5969105
; GENERAL INFORMATION:
; APPLICANT: Mc Wherter, Charles
; APPLICANT: Feng, Yiqing
; TITLE OF INVENTION: No. 5969105el Stem Cell Factor Receptor
; TITLE OF INVENTION: Agonists
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: G. D. Searle & Co.
; STREET: P.O. Box 5110
; CITY: Chicago
; STATE: IL
; COUNTRY: U. S. A.
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,848A
; FILING DATE: 21-OCT-1997
; CLASSIFICATION: 536

; NAME/KEY: misc_feature
; LOCATION: 376
; OTHER INFORMATION: n=a, g, c o r t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 382
; OTHER INFORMATION: s=g o r c
US-09-513-999C-31455

PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/029,165
; FILING DATE: 25-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2992/1
; TELEPHONE: 314-737-6986
; TELEFAX: 314-737-6972
; TELEX:
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-955-848A-46

Query Match      1.4%; Score 18; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 ATCCACAAATGACTTGGC 428
DB 134 ATCCACAAATGACTTGGC 117

RESULT 46
PCT-US95-03866-1/c
; Sequence 1, Application PC/TUS9503866
; GENERAL INFORMATION:
; APPLICANT: CytoMed, Inc. (all states except US)
; APPLICANT: Nocka, Karl (US only)
; APPLICANT: Lobell, Robert B (US only)
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03866
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,379
; FILING DATE: 28-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: CytoMed/2
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```



;; FEATURE: CDS  
;; NAME/KEY: 1..498  
;; LOCATION: 1..498  
;; OTHER INFORMATION: /product= "soluble human kit ligand  
;; OTHER INFORMATION: (amino acids 1-165)"  
PCT-US95-03866-1

Query Match 1.4%; Score 18; DB 5; Length 495;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 ATCCAAACAAATGACTTGGC 428  
Db 134 ATCCAAACAAATGACTTGGC 117

RESULT 47  
PCT-US95-03866-17/c  
; Sequence 17, Application PC/TUS9503866  
; GENERAL INFORMATION:  
; APPLICANT: CytoMed, Inc. (all states except US)  
; APPLICANT: Nocka, Karl (US only)  
; APPLICANT: Lobell, Robert B (US only)  
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND  
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/03866  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/220,379  
; FILING DATE: 28-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr, James F  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: CytoMed/2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 498 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..498  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 4..498  
; OTHER INFORMATION: /product= "human KL w/Tyr->Cys  
; OTHER INFORMATION: substitution at aa 26"  
PCT-US95-03866-17

Query Match 1.4%; Score 18; DB 5; Length 498;  
Best Local Similarity 100.0%; Pred. No. 71;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 ATCCAAACAAATGACTTGGC 428  
Db 137 ATCCAAACAAATGACTTGGC 120

RESULT 48  
PCT-US95-03866-35/c  
; Sequence 35, Application PC/TUS9503866  
; GENERAL INFORMATION:  
; APPLICANT: CytoMed, Inc. (all states except US)  
; APPLICANT: Nocka, Karl (US only)  
; APPLICANT: Lobell, Robert B (US only)  
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND  
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/03866  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/220,379  
; FILING DATE: 28-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haley Jr, James F  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: CytoMed/2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 498 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..498  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 4..498  
; OTHER INFORMATION: /product= "human KL with Cys->Ser  
; OTHER INFORMATION: substitution at aa 138"  
PCT-US95-03866-35

Query Match 1.4%; Score 18; DB 5; Length 498;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 ATCCAAACAAATGACTTGGC 428  
Db 137 ATCCAAACAAATGACTTGGC 120

RESULT 49  
PCT-US95-03866-19/c

```
; Sequence 19, Application PC/TUS9503866
; GENERAL INFORMATION:
; APPLICANT: CytoMed, Inc. (all states except US)
; APPLICANT: Nocka, Karl (US only)
; APPLICANT: Lobell, Robert B (US only)
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,379
; FILING DATE: 28-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: Cytomed/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..501
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 4..501
; OTHER INFORMATION: /product= "human KL w/extra Cys
; OTHER INFORMATION: inserted at aa 27"
; PCT-US95-03866-19

Query Match 1.4%; Score 18; DB 5; Length 501;
Best Local Similarity 100.0%; Pred.No.71;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 ATCCAAACATGACTTGGC 428
DB 140 ATCCAAACATGACTTGGC 123

RESULT 50
US-08-955-848A-58/c
; Sequence 58, Application US/08955848A
; Patent No. 5969105
; GENERAL INFORMATION:
; APPLICANT: Mc Whorter, Charles
; APPLICANT: Feng, Yiding
; TITLE OF INVENTION: No. 5969105el Stem Cell Factor Receptor
; TITLE OF INVENTION: Agonists
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: G. D. Seattle & Co.
; STREET: P.O. Box 5110
; CITY: Chicago
; STATE: IL
; COUNTRY: U. S. A.
; ZIP: 60680
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,848A
; FILING DATE: 21-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/029,165
; FILING DATE: 25-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Dennis A
; REGISTRATION NUMBER: 34,547
; REFERENCE/DOCKET NUMBER: C-2992/1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-737-6986
; TELEFAX: 314-737-6972
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-955-848A-58

Query Match 1.4%; Score 18; DB 2; Length 506;
Best Local Similarity 100.0%; Pred.No.71;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 ATCCAAACATGACTTGGC 428
DB 448 ATCCAAACATGACTTGGC 431

Search completed: April 25, 2005, 03:59:34
Job time : 305 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 02:39:52 ; Search time 2329 Seconds.

(without alignments)

3320.962 Million cell updates/sec

Title: US-10-679-246-1

Perfect score: 1274

Sequence: 1 ttctttagttagttagt.....ctctttagttagttagtgaagc 1274

Scoring table: OLIGO NUC

Gapop\_60.0 , Gapext 60.0

Searched: 5633728 seqs, 3035525691 residues

Word size : 15

Total number of hits satisfying chosen parameters: 17974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published Applications\_NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1274	100.0	1274	18 US-10-679-246-1	Sequence 1, Appli
2	1274	100.0	2924	18 US-10-357-930-22959	Sequence 22959, A
3	1274	100.0	2924	18 US-10-357-930-23457	Sequence 23457, A
4	1274	100.0	2924	18 US-10-357-930-28822	Sequence 28822, A
5	1274	100.0	6107	17 US-09-764-891-8070	Sequence 8070, Ap
6	1223	96.0	2829	17 US-10-108-260A-269	Sequence 269, App
7	1041	81.7	2440	17 US-10-641-643-588	Sequence 588, App
8	356	27.9	1535	9 US-09-925-297-84	Sequence 84, Appli
9	159	12.5	405	18 US-10-357-930-38440	Sequence 38440, A
10	154	12.1	277	18 US-10-357-930-8510	Sequence 8510, Ap
11	139	10.9	417	18 US-10-357-930-43808	Sequence 43808, A

128	10.0	473	18	US-10-357-930-48480	Sequence 48480, A
120	9.4	440	9	US-09-783-590-4667	Sequence 4667, Ap
118	9.3	411	18	US-10-357-930-13819	Sequence 13819, A
118	9.3	441	18	US-10-357-930-34958	Sequence 34958, A
85	6.7	438	18	US-10-357-930-4650	Sequence 4650, Ap
62	4.9	360	9	US-09-960-352-10011	Sequence 10011, A
52	4.1	426	18	US-10-357-930-18669	Sequence 18669, A
19	3.6	188	18	US-10-425-115-63145	Sequence 63145, A
33	2.6	222	10	US-09-764-891-1549	Sequence 1549, Ap
21	2.5	25	19	US-10-719-900-612660	Sequence 612660, A
24	1.9	33	19	US-10-490-581-10	Sequence 10, Appl
23	1.8	226	9	US-09-728-445-159	Sequence 159, App
22	1.7	2466	17	US-10-424-599-69723	Sequence 69723, A
21	1.6	21	19	US-10-490-581-6	Sequence 5, Appli
21	1.6	21	19	US-10-490-581-6	Sequence 6, Appli
21	1.6	25	19	US-10-719-900-37747	Sequence 37747, A
21	1.6	25	19	US-10-719-900-257078	Sequence 257078, A
21	1.6	442	10	US-09-918-995-15609	Sequence 15609, A
21	1.6	442	18	US-10-425-115-119327	Sequence 119327, A
21	1.6	477	10	US-09-918-995-15341	Sequence 15341, A
21	1.6	1094	17	US-10-627-757-10	Sequence 10, Appl
21	1.6	1150	17	US-10-627-757-9	Sequence 9, Appli
21	1.6	46951	16	US-10-091-281-2	Sequence 2, Appli
20	1.6	28	18	US-10-679-246-42	Sequence 42, Appli
20	1.6	257	17	US-10-305-720-397	Sequence 397, App
20	1.6	335	17	US-10-242-535A-55073	Sequence 55073, A
20	1.6	335	17	US-10-085-783A-55073	Sequence 55073, A
20	1.6	497	13	US-10-027-632-317039	Sequence 317039, A
20	1.6	497	13	US-10-027-632-317040	Sequence 317040, A
20	1.6	497	17	US-10-027-632-317039	Sequence 317039, A
20	1.6	497	13	US-10-027-632-317040	Sequence 317040, A
20	1.6	498	13	US-10-027-632-88621	Sequence 88621, A
20	1.6	498	13	US-10-027-632-88622	Sequence 88622, A
20	1.6	498	17	US-10-027-632-88621	Sequence 88621, A
20	1.6	498	17	US-10-027-632-88622	Sequence 88622, A
20	1.6	507	13	US-10-027-632-91273	Sequence 91273, A
20	1.6	507	13	US-10-027-632-317676	Sequence 317676, A
20	1.6	507	17	US-10-027-632-317676	Sequence 317676, A
20	1.6	975	17	US-10-172-118-1117	Sequence 1117, Ap
20	1.6	975	17	US-10-342-887-1117	Sequence 1117, Ap
20	1.6	1182	13	US-10-027-632-249645	Sequence 249645, A
20	1.6	1182	17	US-10-027-632-249645	Sequence 249645, A
20	1.6	2240	15	US-10-177-293-424	Sequence 424, App
20	1.6	2240	18	US-10-717-597-137	Sequence 137, App
20	1.6	2240	18	US-10-755-889-123	Sequence 123, App
20	1.6	2502	15	US-10-177-293-426	Sequence 426, App
20	1.6	2826	18	US-10-425-115-19946	Sequence 19946, A
20	1.6	4329	18	US-10-425-115-19959	Sequence 19959, A
20	1.6	4809	18	US-10-425-115-19949	Sequence 19949, A
20	1.6	3673778	16	US-10-312-841-1	Sequence 1, Appli
19	1.5	19	18	US-10-679-246-43	Sequence 43, Appl
19	1.5	231	18	US-10-357-930-1151	Sequence 1151, Ap
19	1.5	406	18	US-10-425-115-164311	Sequence 164311, A
19	1.5	447	18	US-10-357-930-31518	Sequence 31518, A
19	1.5	447	18	US-10-357-930-40488	Sequence 40488, A
19	1.5	470	18	US-10-357-930-47061	Sequence 47061, A
19	1.5	526	18	US-10-425-115-109691	Sequence 109691, A
19	1.5	564	16	US-10-023-386-12572	Sequence 12572, A
19	1.5	567	18	US-10-021-323-17465	Sequence 17465, A
19	1.5	605	10	US-09-764-891-6710	Sequence 6710, Ap
19	1.5	605	10	US-09-764-891-6710	Sequence 569, App
19	1.5	605	14	US-10-091-572-569	Sequence 571, App
19	1.5	605	14	US-10-091-572-571	Sequence 72924, A
19	1.5	761	17	US-10-424-599-72924	Sequence 10320, A
19	1.5	850	18	US-10-357-930-10320	Sequence 762, App
19	1.5	978	19	US-10-774-355A-762	Sequence 97, Appl
19	1.5	1015	14	US-10-023-601-97	Sequence 33999, A
19	1.5	1264	17	US-10-424-599-33999	Sequence 22304, A
19	1.5	1866	17	US-10-282-122A-22304	Sequence 6, Appli
19	1.5	1884	17	US-10-381-779-6	Sequence 1055, Ap
19	1.5	1995	17	US-10-641-643-1055	Sequence 760, App
19	1.5	1995	19	US-10-741-600-760	





377 17 1.3 201 18 US-10-741-601-24917 Sequence 24917, A 450 17 1.3 429 9 US-09-867-701-8635 Sequence 8635, Ap  
378 17 1.3 201 18 US-10-719-933-13972 Sequence 13972, A 451 17 1.3 429 17 US-10-424-599-22560 Sequence 22560, A  
379 17 1.3 201 18 US-10-719-933-22123 Sequence 22123, A 452 17 1.3 431 9 US-09-983-965-3592 Sequence 3592, Ap  
380 17 1.3 201 18 US-10-719-933-22124 Sequence 22124, A 453 17 1.3 431 13 US-10-027-632-278290 Sequence 278290, Ap  
381 17 1.3 201 18 US-10-719-933-36245 Sequence 36245, A 454 17 1.3 431 17 US-10-027-632-278290 Sequence 278290, Ap  
382 17 1.3 201 18 US-10-719-933-37062 Sequence 37062, A 455 17 1.3 434 17 US-10-027-632-278290 Sequence 278290, Ap  
383 17 1.3 201 18 US-10-719-933-44748 Sequence 44748, A 456 17 1.3 434 9 US-09-850-716A-26 Sequence 26, Appl  
384 17 1.3 201 18 US-10-719-933-50149 Sequence 50149, A 457 17 1.3 434 9 US-09-897-778-26 Sequence 26, Appl  
385 17 1.3 201 18 US-10-719-933-50150 Sequence 50150, A 458 17 1.3 434 14 US-09-466-396A-26 Sequence 26, Appl  
386 17 1.3 201 19 US-10-741-600-9430 Sequence 9430, Ap 459 17 1.3 434 15 US-10-117-982-26 Sequence 26, Appl  
387 17 1.3 201 19 US-10-741-600-34549 Sequence 34549, A 460 17 1.3 434 15 US-10-117-982-26 Sequence 26, Appl  
388 17 1.3 201 19 US-10-741-600-35052 Sequence 35052, A 461 17 1.3 434 17 US-10-313-986-26 Sequence 26, Appl  
389 17 1.3 201 19 US-10-741-600-41956 Sequence 41956, A 462 17 1.3 434 18 US-10-775-972-26 Sequence 26, Appl  
390 17 1.3 201 19 US-10-741-600-41957 Sequence 41957, A 463 17 1.3 436 9 US-09-867-701-2032 Sequence 2032, Ap  
391 17 1.3 201 19 US-10-741-600-52117 Sequence 52117, A 464 17 1.3 437 9 US-09-735-705-74 Sequence 74, Appl  
392 17 1.3 201 19 US-10-741-600-52119 Sequence 52119, A 465 17 1.3 437 9 US-09-850-716A-74 Sequence 74, Appl  
393 17 1.3 201 19 US-10-741-600-52122 Sequence 52122, A 466 17 1.3 437 9 US-09-897-778-74 Sequence 74, Appl  
394 17 1.3 201 19 US-10-741-600-67297 Sequence 67297, A 467 17 1.3 437 10 US-09-466-396A-74 Sequence 74, Appl  
395 17 1.3 201 19 US-10-741-600-67298 Sequence 67298, A 468 17 1.3 437 14 US-10-007-700-74 Sequence 74, Appl  
396 17 1.3 201 19 US-10-674-124A-14204 Sequence 14204, A 469 17 1.3 437 15 US-10-117-982-74 Sequence 74, Appl  
397 17 1.3 221 9 US-09-983-965-5845 Sequence 5845, Ap 470 17 1.3 437 17 US-10-313-986-74 Sequence 74, Appl  
398 17 1.3 222 9 US-09-933-797-671 Sequence 671, Ap 471 17 1.3 437 18 US-10-775-972-74 Sequence 74, Appl  
399 17 1.3 232 9 US-09-960-352-4577 Sequence 4577, Ap 472 17 1.3 437 18 US-10-357-930-13134 Sequence 13134, A  
400 17 1.3 233 17 US-10-242-535A-57047 Sequence 57047, A 473 17 1.3 438 13 US-10-040-739-575 Sequence 575, App  
401 17 1.3 233 17 US-10-085-783A-57047 Sequence 57047, A 474 17 1.3 438 13 US-09-814-353-21978 Sequence 21978, A  
402 17 1.3 251 17 US-10-242-535A-9435 Sequence 9435, Ap 475 17 1.3 444 10 US-10-027-632-74611 Sequence 74611, A  
403 17 1.3 251 17 US-10-085-783A-9435 Sequence 9435, Ap 476 17 1.3 447 13 US-10-027-632-299562 Sequence 299562, A  
404 17 1.3 255 17 US-10-424-599-9069 Sequence 9069, Ap 477 17 1.3 447 13 US-10-027-632-299562 Sequence 299562, A  
405 17 1.3 261 10 US-09-237-183A-2497 Sequence 2497, Ap 478 17 1.3 447 17 US-10-027-632-299562 Sequence 299562, A  
406 17 1.3 267 18 US-10-437-963-155 Sequence 155, App 479 17 1.3 451 18 US-10-357-930-48932 Sequence 48932, A  
407 17 1.3 279 18 US-10-674-124A-23007 Sequence 23007, A 480 17 1.3 451 18 US-10-357-930-48932 Sequence 48932, A  
408 17 1.3 302 17 US-10-242-535A-56462 Sequence 56462, A 481 17 1.3 464 13 US-10-027-632-66783 Sequence 66783, A  
409 17 1.3 302 17 US-10-085-783A-56462 Sequence 56462, A 482 17 1.3 464 13 US-10-027-632-66783 Sequence 66783, A  
410 17 1.3 308 17 US-10-424-599-100302 Sequence 100302, A 483 17 1.3 464 17 US-10-027-632-66783 Sequence 66783, A  
411 17 1.3 311 17 US-10-085-783A-58595 Sequence 58595, A 484 17 1.3 464 17 US-10-027-632-66783 Sequence 66783, A  
412 17 1.3 311 17 US-10-085-783A-58595 Sequence 58595, A 485 17 1.3 464 17 US-10-027-632-66783 Sequence 66783, A  
413 17 1.3 318 18 US-10-437-963-50449 Sequence 50449, A 486 17 1.3 464 17 US-10-027-632-66783 Sequence 66783, A  
414 17 1.3 322 18 US-10-425-115-162462 Sequence 162462, A 487 17 1.3 466 18 US-10-357-930-34281 Sequence 34281, A  
415 17 1.3 323 17 US-10-085-783A-37619 Sequence 37619, A 488 17 1.3 466 18 US-10-357-930-34281 Sequence 34281, A  
416 17 1.3 323 17 US-10-085-783A-37619 Sequence 37619, A 489 17 1.3 466 9 US-09-864-761-9 Sequence 9, Appl  
417 17 1.3 326 9 US-09-960-352-1631 Sequence 1631, Ap 490 17 1.3 469 18 US-10-437-963-5372 Sequence 5372, Ap  
418 17 1.3 326 18 US-10-425-115-35986 Sequence 35986, A 491 17 1.3 470 13 US-10-027-632-35796 Sequence 35796, A  
419 17 1.3 331 17 US-10-242-535A-32558 Sequence 32558, A 492 17 1.3 470 17 US-10-027-632-35796 Sequence 35796, A  
420 17 1.3 331 17 US-10-085-783A-32558 Sequence 32558, A 493 17 1.3 470 17 US-10-424-599-23588 Sequence 23588, A  
421 17 1.3 335 18 US-10-425-115-37944 Sequence 37944, A 494 17 1.3 478 9 US-09-960-352-14432 Sequence 14432, A  
422 17 1.3 339 17 US-10-424-599-138100 Sequence 138100, A 495 17 1.3 479 9 US-09-960-352-14432 Sequence 14432, A  
423 17 1.3 348 17 US-10-149-736-31 Sequence 31, Appl 496 17 1.3 480 10 US-09-918-995-3334 Sequence 3334, Ap  
424 17 1.3 348 19 US-10-424-599-25421 Sequence 25421, A 497 17 1.3 480 10 US-09-918-995-3334 Sequence 3334, Ap  
425 17 1.3 356 17 US-10-424-599-48056 Sequence 48056, A 498 17 1.3 480 17 US-10-085-783A-29395 Sequence 29395, A  
426 17 1.3 357 9 US-09-867-701-7378 Sequence 7378, Ap 499 17 1.3 480 17 US-10-085-783A-29395 Sequence 29395, A  
427 17 1.3 363 18 US-10-425-115-147959 Sequence 147959, A 500 17 1.3 480 17 US-10-085-783A-29395 Sequence 29395, A  
428 17 1.3 368 17 US-10-424-599-9639 Sequence 9639, Ap 501 17 1.3 483 10 US-09-918-995-23896 Sequence 23896, A  
429 17 1.3 370 17 US-10-425-114-7414 Sequence 7414, Ap 502 17 1.3 489 10 US-09-918-995-23896 Sequence 23896, A  
430 17 1.3 370 17 US-10-425-114-7414 Sequence 7414, Ap 503 17 1.3 489 17 US-10-276-774-644 Sequence 644, App  
431 17 1.3 373 10 US-09-814-353-18991 Sequence 18991, A 504 17 1.3 489 17 US-10-276-774-644 Sequence 644, App  
432 17 1.3 376 9 US-09-974-300-8213 Sequence 8213, Ap 505 17 1.3 490 14 US-10-052-283-238 Sequence 238, App  
433 17 1.3 376 9 US-09-974-300-8213 Sequence 8213, Ap 506 17 1.3 492 17 US-10-062-674-365 Sequence 365, App  
434 17 1.3 376 10 US-09-803-719-2074 Sequence 2074, Ap 507 17 1.3 494 9 US-09-864-761-4394 Sequence 4394, Ap  
435 17 1.3 378 18 US-10-425-115-67094 Sequence 67094, A 508 17 1.3 494 18 US-10-424-599-18680 Sequence 18680, A  
436 17 1.3 387 18 US-10-357-930-13230 Sequence 13230, A 509 17 1.3 498 18 US-10-424-599-18680 Sequence 18680, A  
437 17 1.3 408 18 US-10-357-930-19126 Sequence 19126, A 510 17 1.3 500 14 US-10-052-283-214 Sequence 214, App  
438 17 1.3 413 13 US-10-027-632-182883 Sequence 182883, A 511 17 1.3 504 13 US-10-052-283-214 Sequence 214, App  
439 17 1.3 413 13 US-10-027-632-182883 Sequence 182883, A 512 17 1.3 504 13 US-10-052-283-214 Sequence 214, App  
440 17 1.3 418 10 US-09-918-995-36285 Sequence 36285, A 513 17 1.3 504 13 US-10-052-283-214 Sequence 214, App  
441 17 1.3 418 13 US-10-027-632-297609 Sequence 297609, A 514 13 US-10-027-632-62778 Sequence 62778, A  
442 17 1.3 418 13 US-10-027-632-297609 Sequence 297609, A 515 13 US-10-027-632-62778 Sequence 62778, A  
443 17 1.3 418 17 US-10-027-632-81800 Sequence 81800, A 516 13 US-10-027-632-62778 Sequence 62778, A  
444 17 1.3 418 17 US-10-027-632-81800 Sequence 81800, A 517 13 US-10-027-632-62778 Sequence 62778, A  
445 17 1.3 421 18 US-10-027-632-297609 Sequence 297609, A 518 17 1.3 514 17 US-10-027-632-62778 Sequence 62778, A  
446 17 1.3 421 18 US-10-027-632-297609 Sequence 297609, A 519 17 1.3 514 17 US-10-027-632-62778 Sequence 62778, A  
447 17 1.3 423 9 US-09-960-352-11252 Sequence 11252, A 520 17 1.3 514 17 US-10-027-632-62778 Sequence 62778, A  
448 17 1.3 423 17 US-10-424-599-118097 Sequence 118097, A 521 17 1.3 516 13 US-10-027-632-62778 Sequence 62778, A  
449 17 1.3 424 10 US-09-918-995-17769 Sequence 17769, A 522 17 1.3 516 13 US-10-027-632-62778 Sequence 62778, A  
450 17 1.3 428 18 US-10-357-930-19027 Sequence 19027, A 523 17 1.3 516 13 US-10-027-632-62778 Sequence 62778, A









961 17 1.3 5735 9 US-09-728-952-44 Sequence 44, App1  
962 17 1.3 5735 9 US-09-728-952-99 Sequence 99, App1  
963 17 1.3 5759 15 US-10-311-455-1677 Sequence 1677, Ap  
964 17 1.3 5768 15 US-10-311-455-2135 Sequence 2135, Ap  
965 17 1.3 5807 15 US-10-311-455-1128 Sequence 1128, Ap  
966 17 1.3 5815 17 US-10-257-166-105 Sequence 105, App  
967 17 1.3 5823 15 US-10-240-453-256 Sequence 256, App  
968 17 1.3 5890 15 US-10-311-455-1311 Sequence 1311, Ap  
969 17 1.3 5971 15 US-10-311-455-356 Sequence 356, App  
970 17 1.3 6014 17 US-10-221-613-367 Sequence 367, App  
971 17 1.3 6014 17 US-10-221-613-367 Sequence 367, App  
972 17 1.3 6014 17 US-10-221-613-368 Sequence 368, App  
973 17 1.3 6024 15 US-10-311-455-1619 Sequence 1619, Ap  
974 17 1.3 6042 15 US-10-311-455-1917 Sequence 1917, Ap  
975 17 1.3 6049 15 US-10-311-455-346 Sequence 346, App  
976 17 1.3 6061 14 US-10-239-676-40 Sequence 40, App1  
977 17 1.3 6062 15 US-10-311-455-2051 Sequence 2051, Ap  
978 17 1.3 6074 15 US-10-311-455-1037 Sequence 1037, Ap  
979 17 1.3 6081 15 US-10-311-455-1401 Sequence 1401, Ap  
980 17 1.3 6092 15 US-10-311-455-993 Sequence 993, App  
981 17 1.3 6109 15 US-10-311-455-1704 Sequence 1704, Ap  
982 17 1.3 6112 15 US-10-311-455-445 Sequence 445, App  
983 17 1.3 6113 15 US-10-311-455-776 Sequence 776, App  
984 17 1.3 6131 15 US-10-311-455-863 Sequence 863, App  
985 17 1.3 6145 17 US-10-221-714A-269 Sequence 269, App  
986 17 1.3 6171 15 US-10-311-455-983 Sequence 983, App  
987 17 1.3 6182 15 US-10-311-455-1987 Sequence 1987, Ap  
988 17 1.3 6197 15 US-10-311-455-1683 Sequence 1683, Ap  
989 17 1.3 6204 15 US-10-311-455-1800 Sequence 1800, Ap  
990 17 1.3 6246 15 US-10-172-086-27 Sequence 27, App1  
991 17 1.3 6246 15 US-10-311-455-989 Sequence 989, App  
992 17 1.3 6246 18 US-10-311-507-49 Sequence 49, App1  
993 17 1.3 6246 18 US-10-480-846-27 Sequence 27, App1  
994 17 1.3 6254 15 US-10-311-455-1593 Sequence 1593, Ap  
995 17 1.3 6298 14 US-10-239-676-63 Sequence 63, App1  
996 17 1.3 6298 15 US-10-240-453-67 Sequence 67, App1  
997 17 1.3 6303 18 US-10-433-793-115 Sequence 115, App  
998 17 1.3 6306 14 US-10-239-676-223 Sequence 223, App  
999 17 1.3 6306 15 US-10-240-453-331 Sequence 331, App  
c1000 17 1.3 6309 15 US-10-311-455-277 Sequence 277, App

ALIGNMENTS

RESULT 1  
US-10-679-246-1  
; Sequence 1, Application US/10679246  
; Publication No. US20040163138A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Matsuzawa, Shu-ichi  
; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved  
; TITLE OF INVENTION: in Protein Degradation, Products and Methods Related Thereto  
; FILE REFERENCE: 66821-235  
; CURRENT APPLICATION NUMBER: US/10/679,246  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: US 09/591,694  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1274  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (274)...(1167)  
US-10-679-246-1

Query Match 100.0%; Score 1274; DB 18; Length 1274;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTTCTTTAGTTGTTTATGGTCCATTTTCTATTATTAGCAATTTATTATTCTATGTAGTCAT 60  
1 TTTCTTTAGTTGTTTATGGTCCATTTTCTATTATTAGCAATTTATTATTCTATGTAGTCAT 60  
61 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAAACATTTTGAAGAGAGCGTTATC 120  
61 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAAACATTTTGAAGAGAGCGTTATC 120  
121 CAGTGTACAGATCCTTAATAAAGTGACATTCAGTGTATAATTTTATTATTATTATTATCTTTT 180  
121 CAGTGTACAGATCCTTAATAAAGTGACATTCAGTGTATAATTTTATTATTATTATTATCTTTT 180  
181 TTAATCTTATTTTCTTCT 240  
181 TTAATCTTATTTTCTTCT 240  
241 TATGCGATGTAACACATTTATTTATAAGTAAGTCAATGTTTATTTTCTCTCTCTCTCTCTCT 300  
241 TATGCGATGTAACACATTTATTTATAAGTAAGTCAATGTTTATTTTCTCTCTCTCTCTCTCT 300  
301 CCTTATGTTATTTTATTTTATGAAATGAGCGTCTGAGTGTCTAGCAATTAACCTACCGGTACC 360  
301 CCTTATGTTATTTTATTTTATGAAATGAGCGTCTGAGTGTCTAGCAATTAACCTACCGGTACC 360  
361 TCGAAGTGTCCACCATCCAGAGGGTCTGCTGCTGAGTGTGAGCAATCGCATCCAACAAAT 420  
361 TCGAAGTGTCCACCATCCAGAGGGTCTGCTGCTGAGTGTGAGCAATCGCATCCAACAAAT 420  
421 GACTTGGCGAGTCTTTTGGGATCCATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTC 480  
421 GACTTGGCGAGTCTTTTGGGATCCATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTC 480  
481 CAATGTACAGATGGGCACTCTTGTGTAGCACTCTGCGCCCAAAAGCTCACATGTTGTCCA 540  
481 CAATGTACAGATGGGCACTCTTGTGTAGCACTCTGCGCCCAAAAGCTCACATGTTGTCCA 540  
541 ACTTCCCGGGGGCCCTTTGGGATCCATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTC 600  
541 ACTTCCCGGGGGCCCTTTGGGATCCATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAATTC 600  
601 GTACTTTTCCCTGTAAATATGCGTCTTCTGTGATGAAATTAACCTCTGCCACACACAGAA 660  
601 GTACTTTTCCCTGTAAATATGCGTCTTCTGTGATGAAATTAACCTCTGCCACACACAGAA 660  
661 AAAGCAGACCATGAAGAGCTCTGTGAGTGTAGGCGCTTATTCCTGTCGTCCTGGTGTCT 720  
661 AAAGCAGACCATGAAGAGCTCTGTGAGTGTAGGCGCTTATTCCTGTCGTCCTGGTGTCT 720  
721 TCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCATAG 780  
721 TCCTGTAAATGGCAAGGCTCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCATAG 780  
781 TCCATTACACCCCTACAGGAGGATATAGTTTTTCTGTACAGACATTAATCTTCTCT 840  
781 TCCATTACACCCCTACAGGAGGATATAGTTTTTCTGTACAGACATTAATCTTCTCTCT 840  
841 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTGGCTTTTCACTTCACTGTTAGTCTTA 900  
841 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTGGCTTTTCACTTCACTGTTAGTCTTA 900  
901 GAGAAACAGGAAAAATACGATGGTCAACGAGTCTTCCGCAATCGTACAGCTGATAGGA 960  
901 GAGAAACAGGAAAAATACGATGGTCAACGAGTCTTCCGCAATCGTACAGCTGATAGGA 960  
961 ACACCGACGACGCTGAAATTTTCTTACCGACTTGGCTTAAATGGTCTATAGGACCA 1020  
961 ACACCGACGACGCTGAAATTTTCTTACCGACTTGGCTTAAATGGTCTATAGGACCA 1020  
1021 TTGACTTTGGGAACGACTCTCTCGATCTATTCATGAAGGAATTCGAACAGCCATTAATGAAT 1080  
1021 TTGACTTTGGGAACGACTCTCTCGATCTATTCATGAAGGAATTCGAACAGCCATTAATGAAT 1080

QY 1081 AGCGACTGTCTAGTCTTTGACACCCAGCATGCGACAGCTTTTTCAGAAAAATGCCAATTTA 1140  
DB 1081 AGCGACTGTCTAGTCTTTGACACCCAGCATGCGACAGCTTTTTCAGAAAAATGCCAATTTA 1140  
QY 1141 GGCAATCAATGTAACATTTTCCATGCTGTCGAAATGGCAATCAAAACATTTTCTGGCCAGTGT 1200  
DB 1141 GGCAATCAATGTAACATTTTCCATGCTGTCGAAATGGCAATCAAAACATTTTCTGGCCAGTGT 1200  
QY 1201 TTAAACCTTCAGTTTTCACAGAAAAATAGGACCCATCTGCTGCGCAACCTTAAACCTCTTTT 1260  
DB 1201 TTAAACCTTCAGTTTTCACAGAAAAATAGGACCCATCTGCTGCGCAACCTTAAACCTCTTTT 1260  
QY 1261 CGGTAGGTGGAAGC 1274  
DB 1261 CGGTAGGTGGAAGC 1274

## RESULT 2

US-10-357-930-22959  
; Sequence 22959, Application US/10357930  
; Publication No. US20040259086A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Endege, Wilson  
; APPLICANT: Monahan, John  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
; FILE REFERENCE: MRI-007BCN  
; CURRENT APPLICATION NUMBER: US/10/357,930  
; CURRENT FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: 09/785,276  
; PRIOR FILING DATE: 2003-02-16  
; PRIOR APPLICATION NUMBER: 60/183,319  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: 60/189,862  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/207,454  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/211,314  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/219,007  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/255,281  
; PRIOR FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 62232  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22959  
; LENGTH: 2924  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1..57, 2923, 2924  
; OTHER INFORMATION: n = A,T,C or G  
US-10-357-930-22959

Query Match 100.0%; Score 1274; DB 18; Length 2924;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTTAGTGTGTTATGGTCCATTTTCTATTATTTAGCATTTATTTCTATGTAGTCTAT 60  
DB 480 TTTCTTTAGTGTGTTATGGTCCATTTTCTATTATTTAGCATTTATTTCTATGTAGTCTAT 539  
QY 61 CCAAGACGATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120  
DB 540 CCAAGACGATTAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 599  
QY 121 CAGTGTACAGATCCTAAATAGGACGACATTCAGTGTAAATTTTATTTTATATCTTTT 180  
DB 600 CAGTGTACAGATCCTAAATAGGACGACATTCAGTGTAAATTTTATTTTATATCTTTT 659

QY 181 TTAATCCCTATTTTCTTCTCTCTTTGCTCAGTAAATTTTGTATGAAACCTTTAAAGGACT 240  
DB 660 TTAATCCCTATTTTCTTCTCTCTTTGCTCAGTAAATTTTGTATGAAACCTTTAAAGGACT 719  
QY 241 TATGGCATGTAACATTTATTTATAAGTAAGTCAATGTTTATATTTTCTCTGCGCT 300  
DB 720 TATGGCATGTAACATTTATTTATAAGTAAGTCAATGTTTATATTTTCTCTGCGCT 779  
QY 301 CTTTATGTTATTTTATTTTTCAGAAATGAGCCGTGAGCTGCTACAGCATTTACCTACCGGTACC 360  
DB 780 CTTTATGTTATTTTATTTTTCAGAAATGAGCCGTGAGCTGCTACAGCATTTACCTACCGGTACC 839  
QY 361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCGCTGAGCTGCGACAACTGATGCAACAAT 420  
DB 840 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCGCTGAGCTGCGACAACTGATGCAACAAT 899  
QY 421 GACTTGGCGAGTCTTTTTCAGTGTCTCAGTCTGCTTTGAGCTATGCTGTACCGCCATCTTT 480  
DB 900 GACTTGGCGAGTCTTTTTCAGTGTCTCAGTCTGCTTTGAGCTATGCTGTACCGCCATCTTT 959  
QY 481 CAATGTGAGAGTGGCCATCTTGTGTTGTAGCAACTGTCGCCCAAAAGCTCAATGTTGTCCA 540  
DB 960 CAATGTGAGAGTGGCCATCTTGTGTTGTAGCAACTGTCGCCCAAAAGCTCAATGTTGTCCA 1019  
QY 541 ACTTGGCGGGGCCCTTTTGGGATCCATTTGCAACTGTCGCCCAAAAGCTCAATGTTGTCCA 600  
DB 1020 ACTTGGCGGGGCCCTTTTGGGATCCATTTGCAACTGTCGCCCAAAAGCTCAATGTTGTCCA 1079  
QY 601 GTACTTTTCCCTGTAAATATATGCGTCTCTGAGTGTGAAATTAACCTGCGCACACAGAA 660  
DB 1080 GTACTTTTCCCTGTAAATATATGCGTCTCTGAGTGTGAAATTAACCTGCGCACACAGAA 1139  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTGTAGGCGCTTATTCCTGTCGCGCCCTGCTGCT 720  
DB 1140 AAAGCAGACCATGAAGAGCTCTGTGAGTGTAGGCGCTTATTCCTGTCGCGCCCTGCTGCT 1199  
QY 721 TCCTGTAATGGAAGGCTCTCTGAGTGTGTAATGCCCATCTGATGATGCAATGCAATAG 780  
DB 1200 TCCTGTAATGGAAGGCTCTCTGAGTGTGTAATGCCCATCTGATGATGCAATGCAATAG 1259  
QY 781 TCCATTACAAACCTACAGGAGAGATATAGTTTTTCTTGTACAGACATTAATCTTCTCT 840  
DB 1260 TCCATTACAAACCTACAGGAGAGATATAGTTTTTCTTGTACAGACATTAATCTTCTCT 1319  
QY 841 GGTGCTGTTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTTCACTTCATGTTAGTCTTA 900  
DB 1320 GGTGCTGTTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTTCACTTCATGTTAGTCTTA 1379  
QY 901 GAGAAACAGGAAAAATACGATGCTCAGAGCTCTTCGCAATCGTACAGCTGATAGGA 960  
DB 1380 GAGAAACAGGAAAAATACGATGCTCAGAGCTCTTCGCAATCGTACAGCTGATAGGA 1439  
QY 961 ACACCGAAGCAGCTGAAATTTTCTTACCCGCTTGAAGTAAATGCTCATAGGCGACGA 1020  
DB 1440 ACACCGAAGCAGCTGAAATTTTCTTACCCGCTTGAAGTAAATGCTCATAGGCGACGA 1499  
QY 1021 TTGACTTGGGAAGCAGCTCTCTCGATCTATTTCATGAAGGAATTTGCAACAGCCATTTAAT 1080  
DB 1500 TTGACTTGGGAAGCAGCTCTCTCGATCTATTTCATGAAGGAATTTGCAACAGCCATTTAAT 1559  
QY 1081 AGCGATGCTGATGCTTTGACACCGAGCATTTGACAGCTTTTTCAGAGAAATGCAATTTA 1140  
DB 1560 AGCGATGCTGATGCTTTGACACCGAGCATTTGACAGCTTTTTCAGAGAAATGCAATTTA 1619  
QY 1141 GGCATCAATGTAACATTTTCCATGTTGAAATGCAATCAAACTTTTCTGCCAGTCT 1200  
DB 1620 GGCATCAATGTAACATTTTCCATGTTGAAATGCAATCAAACTTTTCTGCCAGTCT 1679  
QY 1201 TTAACCTTCAGTTTTCAGAGAAAAATAGGACCCATCTGCTGCGCAACCTTAAACCTTTT 1260  
DB 1680 TTAACCTTCAGTTTTCAGAGAAAAATAGGACCCATCTGCTGCGCAACCTTAAACCTTTT 1739  
QY 1261 CGGTAGGTGGAAGC 1274

[illegible]

```

RESULT 4
; US-10-357-930-28822
; Sequence 28822, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

```



Query Match	100.0%	Score 1274;	DB 18;	Length 2924;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1274;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	TTTCTTTAGTGTGTTATGGTCCATTTCTATTTTAGCATTTATTTATCTCATGTAGTCTAT	60
DB	480	TTTCTTTAGTGTGTTATGGTCCATTTCTATTTTAGCATTTATTTATCTCATGTAGTCTAT	539
QY	.61	CCAAAGACGATTAAAGGAGTTCACATGTTTTTCGGAAACATTTTTCGAAAGAGCTTATC	120
DB	540	CCAAAGACGATTAAAGGAGTTCACATGTTTTTCGGAAACATTTTTCGAAAGAGCTTATC	599
QY	1.21	CAGTGTACAGATCCTAATAAGTGCACATTCAGTGTAAATTTATTTTAAATATCTTTT	180
DB	600	CAGTGTACAGATCCTAATAAGTGCACATTCAGTGTAAATTTATTTTAAATATCTTTT	659
QY	1.81	TTAATCCTATTTTCTTCCCTTTTTGTCACGTAAATTTGTATGAACCTTTAAAGGACT	240
DB	660	TTAATCCTATTTTCTTCCCTTTTTGTCACGTAAATTTGTATGAACCTTTAAAGGACT	719
QY	2.41	TATGGCATGTAAACATTTATTTATAAGTACGTATGTTTAAATTTATTTTCTCTCGCT	300
DB	720	TATGGCATGTAAACATTTATTTATAAGTACGTATGTTTAAATTTATTTTCTCTCGCT	779
QY	3.01	CCTTATGTATTTATTTACAGAAATGAGCGCTCAGACTGCTACAGCAATTACCTACCGGTACC	360
DB	780	CCTTATGTATTTATTTACAGAAATGAGCGCTCAGACTGCTACAGCAATTACCTACCGGTACC	839
QY	3.61	TCGAAGTGTCCACCATCCAGAGGGTCTGCCCTGACTGCGACAACTGCATCCAACAAT	420
DB	840	TCGAAGTGTCCACCATCCAGAGGGTCTGCCCTGACTGCGACAACTGCATCCAACAAT	899
QY	4.21	GACTTGCGAGTCTTTTTTGATGTCCAGTCTGCTTTTGACTATGTGTACCGCCCATCTTT	480
DB	900	GACTTGCGAGTCTTTTTTGATGTCCAGTCTGCTTTTGACTATGTGTACCGCCCATCTTT	959
QY	4.81	CAATGTACAGATGSCCATCTTTGTTGTAGCAACTGTCGCCAAAGCTCACATGTTGTCCA	540
DB	960	CAATGTACAGATGSCCATCTTTGTTGTAGCAACTGTCGCCAAAGCTCACATGTTGTCCA	1019
QY	5.41	ACTTGGCGGGCCCTTTTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGCGCTAATTCA	600



```
QY 1 TTTCTTTAGTGTGTTATGGTCAATTTCTATTTTATAGCAATTTATTTCTATGTAGTCTAT 60
Db 3164 TTTCTTTAGTGTGTTATGGTCAATTTCTATTTTATAGCAATTTATTTCTATGTAGTCTAT 3223
QY 61 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAAACATTTTGAAAGAGAGAGCTTATC 120
Db 3224 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAAACATTTTGAAAGAGAGAGCTTATC 3283
QY 121 CAGTGTACAGATCCTAATAAGTGACACATTCAGTGTAAATTTTATTTTATTTTATTTT 180
Db 3284 CAGTGTACAGATCCTAATAAGTGACACATTCAGTGTAAATTTTATTTTATTTTATTTT 3343
QY 181 TTAATCCTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 3344 TTAATCCTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3403
QY 241 TATGGCATGTAACATTTATTTATAAGTAAGTCATGGTTATATTTATTTTCTCTCGCT 300
Db 3404 TATGGCATGTAACATTTATTTATAAGTAAGTCATGGTTATATTTATTTTCTCTCGCT 3463
QY 301 CTTTATGTATTTATTTTCAAGATAGCCGTGACATGCTACAGCAATTTACCTACCGGTACC 360
Db 3464 CTTTATGTATTTATTTTCAAGATAGCCGTGACATGCTACAGCAATTTACCTACCGGTACC 3523
QY 361 TCGAAGTGTCCACATCCACAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 3524 TCGAAGTGTCCACATCCACAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3583
QY 421 GACTTGGCGAGTCTTTTTCAGTGTCCAGTGTGCTTGTGACTATGTGTTACCGCCCATCTTT 480
Db 3584 GACTTGGCGAGTCTTTTTCAGTGTCCAGTGTGCTTGTGACTATGTGTTACCGCCCATCTTT 3643
QY 481 CAATGTACAGATGGCAATCTTGTGTGTAGCAACTGTGCGCCCAAGCTCAATGTTGTCCA 540
Db 3644 CAATGTACAGATGGCAATCTTGTGTGTAGCAACTGTGCGCCCAAGCTCAATGTTGTCCA 3703
QY 541 ACTTGGCGGGCCCTTGGATGCCATTCGCACTTGGCTATGGAAGAGTGGCTAATTTCA 600
Db 3704 ACTTGGCGGGCCCTTGGATGCCATTCGCACTTGGCTATGGAAGAGTGGCTAATTTCA 3763
QY 601 GTACTTTTCCCTGTAAATATGCGTCTCTGGATGTGAATAACTCTGTCACACACAGAA 660
Db 3764 GTACTTTTCCCTGTAAATATGCGTCTCTGGATGTGAATAACTCTGTCACACACAGAA 3823
QY 661 AAGACGACCAAGACGCTCTGTGAGTTTAGCCCTTATTCCTGTCGGTCCCTGCTGCTGCT 720
Db 3824 AAGACGACCAATGAAGAGCTCTGTGAGTTTAGCCCTTATTCCTGTCGGTCCCTGCTGCT 3883
QY 721 TCCGTAAATGGCAAGCTCTCTGGATGCTGTAAATGCCCATCTGATGCATCAGCATAAG 780
Db 3884 TCCGTAAATGGCAAGCTCTCTGGATGCTGTAAATGCCCATCTGATGCATCAGCATAAG 3943
QY 781 TCCATTACAAACCTACAGGAGAGGATATAGTTTTTCTGTACACAGCAATTAATCTTCCT 840
Db 3944 TCCATTACAAACCTACAGGAGAGGATATAGTTTTTCTGTACACAGCAATTAATCTTCCT 4003
QY 841 GGTGCTGTTGACTGGGTGATGACGAGTCTGTGTTTTGGCTTTTCACTTCACTGTTAGTCTTA 900
Db 4004 GGTGCTGTTGACTGGGTGATGACGAGTCTGTGTTTTGGCTTTTCACTTCACTGTTAGTCTTA 4063
QY 901 GAGAAAACAGGAAAATACGATGTGTACAGAGTCTCTTCGCAATCGTACAGCTGTAGGA 960
Db 4064 GAGAAAACAGGAAAATACGATGTGTACAGAGTCTCTTCGCAATCGTACAGCTGTAGGA 4123
QY 961 ACACGCAAGCAAGCTGAAAAATTTTGTCTTACCGACTTGAGCTTAATGGTCATAGGACGA 1020
Db 4124 ACACGCAAGCAAGCTGAAAAATTTTGTCTTACCGACTTGAGCTTAATGGTCATAGGACGA 4183
QY 1021 TTGACTTGGGAAGCGACTCTTCGATCTATTTCATGAAGGAATTCACAGCCATTATGAAT 1080
Db 4184 TTGACTTGGGAAGCGACTCTTCGATCTATTTCATGAAGGAATTCACAGCCATTATGAAT 4243
```

```
QY 1081 AGCGACTGTCTAGTCTTTTGACACACAGCAATTCACAGCTTTTTCAGAAAAATGGCAATTTA 1140
Db 4244 AGCGACTGTCTAGTCTTTTGACACACAGCAATTCACAGCTTTTTCAGAAAAATGGCAATTTA 4303
QY 1141 GGCATCAATGTAACTATTTTCCATGTTGAAATGGCAATCAAACTTTTCTGCGCCAGTGT 1200
Db 4304 GGCATCAATGTAACTATTTTCCATGTTGAAATGGCAATCAAACTTTTCTGCGCCAGTGT 4363
QY 1201 TTAACACTTCAGTCTTCACAGAAAAATAAGGCAACCATCTGTCTGCAACTAAAACTCTTTT 1260
Db 4364 TTAACACTTCAGTCTTCACAGAAAAATAAGGCAACCATCTGTCTGCAACTAAAACTCTTTT 4423
QY 1261 CGGTAGTGTGAAGC 1274
Db 4424 CGGTAGTGTGAAGC 4437

RESULT 6
US-10-108-260A-269
; Sequence 269, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 269
; LENGTH: 2829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-269

Query Match 96.0%; Score 1223; DB 17; Length 2829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTTCTTTAGTGTGTTATGGTCCATTTTCTATTTTATAGCAATTTATTTCTATGTAGTCTAT 60
Db 1134 TTTCTTTAGTGTGTTATGGTCCATTTTCTATTTTATAGCAATTTATTTCTATGTAGTCTAT 1193
QY 61 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAAACATTTTGAAAGAGAGAGCTTATC 120
Db 1194 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAAACATTTTGAAAGAGAGAGCTTATC 1253
QY 121 CAGTGTACAGATCCTAATAAGTGACACATTCAGTGTAAATTTTATTTTATTTTATTTT 180
Db 1254 CAGTGTACAGATCCTAATAAGTGACACATTCAGTGTAAATTTTATTTTATTTTATTTT 1313
QY 181 TTAATCCTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 1314 TTAATCCTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1373
QY 241 TATGCATGTAAACATTTATTTATAAGTAAGTCATGGTTATATTTTATTTTCTCTGCTCT 300
Db 1374 TATGCATGTAAACATTTATTTATAAGTAAGTCATGGTTATATTTTATTTTCTCTGCTCT 1433
QY 301 CTTTATGTATTTTATTTTCAGAAATGAGCGTACAGCTGTACAGCAATTTACCTACCGGTACC 360
Db 1434 CTTTATGTATTTTATTTTCAGAAATGAGCGTACAGCTGTACAGCAATTTACCTACCGGTACC 1493
QY 361 TCGAAGTGTCCACCATCCACAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 1494 TCGAAGTGTCCACCATCCACAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553
QY 421 GACTTGGCGAGTCTTTTTCAGTGTCCAGTGTCCAGTGTGCTGTTTGTAGTATGTGTACCGCCATCTTT 480
Db 1554 GACTTGGCGAGTCTTTTTCAGTGTCCAGTGTCCAGTGTGCTGTTTGTAGTATGTGTACCGCCATCTTT 1613
QY 481 CAATGTACAGTGGCCATCTTGTGTGTAGCAACTGTGCGCCCAAGCTCAATGTTGTGTCCA 540
```



Db	1007	GCATAAGTCCATTACAACCCCTACAGGGAGAGGATATAGTTTTCTTGCTACAGACATTAA	1066
Qy	834	TCCTCCGGTGCCTGTTGACTCGGGTGATGATGCAGTCTCTGTTTTGGCTTTTCAC TTCATGTT	893
Db	1067	TCCTCCGGTGCCTGTTGACTCGGGTGATGATGCAGTCTCTGTTTTGGCTTTTCAC TTCATGTT	1126
Qy	894	AGTCTTAGAAGAAACAGAGAAAATAAGATGTTGTCACGAGCTTCTTCGCAATCGTACAGCT	953
Db	1127	AGTCTTAGAAGAAACAGAGAAAATAAGATGTTGTCACGAGCTTCTTCGCAATCGTACAGCT	1186
Qy	954	GATAGGAACACGCAAGCAAGCTGAATAATTTTGCTTACCGACTTTGAGCTAAATGGTCATAG	1013
Db	1187	GATAGGAACACGCAAGCAAGCTGAATAATTTTGCTTACCGACTTTGAGCTAAATGGTCATAG	1246
Qy	1014	CGCAGCATTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATTCGAACAGCCAT	1073
Db	1247	CGCAGCATTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATTCGAACAGCCAT	1306
Qy	1074	TATGAATAGCGACTGCTAGTCTTTTGACACGACGATTGCACAGCTTTTTCGAGAAAATGG	1133
Db	1307	TATGAATAGCGACTGCTAGTCTTTTGACACGACGATTGCACAGCTTTTTCGAGAAAATGG	1366
Qy	1134	CAATTTAGGCATCAATGTAACTATTTCCATGTGTTGAAATGGCAATCAAA CATTTTCTGG	1193
Db	1367	CAATTTAGGCATCAATGTAACTATTTCCATGTGTTGAAATGGCAATCAAA CATTTTCTGG	1426
Qy	1194	CCAGTGTTTAAAACTTCAGTTTTCACAGAAAATAGGCACCCTCTGTCTGCCAACCTAAA	1253
Db	1427	CCAGTGTTTAAAACTTCAGTTTTCACAGAAAATAGGCACCCTCTGTCTGCCAACCTAAA	1486
Qy	1254	ACTCTTTTCGGTAGTGGGAAC	1274
Db	1487	ACTCTTTTCGGTAGTGGGAAC	1507

```

RESULT 8
US-09-925-297-84
; Sequence 84, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 84
; LENGTH: 1535
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-297-84

```

Qy	1099	GACACGAGCATTCGACAGCTTTTTTCAGAAAAATGGCAATTTTAGGCATCAATGTAACCTATT	1151
Db	216	GACACGAGCATTCGACAGCTTTTTTCAGAAAAATGGCAATTTTAGGCATCAATGTAACCTATT	275
Qy	1159	TCCATGTGTTGAATAGGCAATCAAAACATTTTCGGCCAGTGTTTAAAACTTCAGTTTCAC	1218
Db	276	TCCATGTGTTGAATAGGCAATCAAAACATTTTCGGCCAGTGTTTAAAACTTCAGTTTCAC	335
Qy	1219	AGAAAAATAAGGCACCCATCTGTCGCAACCTTAAAACTCTTTTCGGTAGTGGGAAGC	1274
Db	336	AGAAAAATAAGGCACCCATCTGTCGCAACCTTAAAACTCTTTTCGGTAGTGGGAAGC	391

```

RESULT 9
US-10-357-930-38440/c
; Sequence 38440, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38440
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-38440

```

RESULT 10  
US-10-357-930-8510/c  
; Sequence 8510, Application US/10357930  
; Publication No. US2004025908A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Endege, Wilson

```
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8510
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 213-214, 216, 218, 220, 224, 225, 233, 260, 265, 269
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-8510

Query Match      12.1%; Score 154; DB 18; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.2e-70;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 TAACTATTTTCTCTCTTTGCTCAGTAAATTTGATGAAACTTTAAAGACTT 241
Db 212 TAACTATTTTCTCTCTTTGCTCAGTAAATTTGATGAAACTTTAAAGACTT 153

QY 242 ATGGCATGTAACATATTTATTAAGTAAATGATGTTATATTTTCTCTGCTC 301
Db 152 ATGGCATGTAACATATTTATTAAGTAAATGATGTTATATTTTCTCTGCTC 93

QY 302 CTTATGATTTATTTTTCAGAAATGAGCCGTCAGAC 335
Db 92 CTTATGATTTATTTTCAGAAATGAGCCGTCAGAC 59

RESULT 11
US-10-357-930-43808/c
; Sequence 43808, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43808
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-43808

Query Match      10.0%; Score 128; DB 18; Length 473;
Best Local Similarity 100.0%; Pred. No. 1.3e-56;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 969 GCAAGCTGAAATTTTGTCTACCGACTTGAGTAAATGTCATAGCGCGACTTG 1028
Db 473 GCAAGCTGAAATTTTGTCTACCGACTTGAGTAAATGTCATAGCGCGACTTG 414
```

```
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43808
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-43808

Query Match      10.9%; Score 139; DB 18; Length 417;
Best Local Similarity 100.0%; Pred. No. 2e-62;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1136 ATTTAGGCATCAATCTTAATTTCCATGCTGTGAAATGGCAATCAAACTTTCTGGCC 1195
Db 413 ATTTAGGCATCAATCTTAATTTCCATGCTGTGAAATGGCAATCAAACTTTCTGGCC 354

QY 1196 AGTGTTTAAAACTTCAGTTTTCACAGAAAATAAGGCACCCATCTGTGCAACCTAAAAC 1255
Db 353 AGTGTTTAAAACTTCAGTTTTCACAGAAAATAAGGCACCCATCTGTGCAACCTAAAAC 294

QY 1256 TCTTTCCGTAGTGGAAGC 1274
Db 293 TCTTTCCGTAGTGGAAGC 275

RESULT 12
US-10-357-930-48480/c
; Sequence 48480, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48480
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-48480

Query Match      10.0%; Score 128; DB 18; Length 473;
Best Local Similarity 100.0%; Pred. No. 1.3e-56;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 969 GCAAGCTGAAATTTTGTCTACCGACTTGAGTAAATGTCATAGCGCGACTTG 1028
Db 473 GCAAGCTGAAATTTTGTCTACCGACTTGAGTAAATGTCATAGCGCGACTTG 414
```



```
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13819
; LENGTH: 411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-13819

Query Match          9.3%; Score 118; DB 18; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.6e-51;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 TTTCATGTTGAAATGGAATCAAAACATTTTCTGCGCCAGTGTTTTAAAACTTCAGTTTC 1216
      |||||||
Db 364 TTTCATGTTGAAATGGAATCAAAACATTTTCTGCGCCAGTGTTTTAAAACTTCAGTTTC 305
      |||||||

QY 1217 ACAGAAATAAGGACCCCATCTCTGTCGCCAACCTAAACTTTTCGGTAGGTGGAAGC 1274
      |||||||
Db 304 ACAGAAATAAGGACCCCATCTCTGTCGCCAACCTAAACTTTTCGGTAGGTGGAAGC 247
      |||||||

RESULT 15
US-10-357-930-34958/c
; Sequence 34958, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34958
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-34958

Query Match          9.3%; Score 118; DB 18; Length 441;
Best Local Similarity 100.0%; Pred. No. 2.6e-51;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 TTTCATGTTGAAATGGAATCAAAACATTTTCTGCGCCAGTGTTTTAAAACTTCAGTTTC 1216
      |||||||
```

```
Db 399 TTTCATGTTGAAATGGAATCAAAACATTTTCTGCGCCAGTGTTTTAAAACTTCAGTTTC 340
QY 1217 ACAGAAATAAGGACCCCATCTCTGTCGCCAACCTAAACTTTTCGGTAGGTGGAAGC 1274
      |||||||
Db 339 ACAGAAATAAGGACCCCATCTCTGTCGCCAACCTAAACTTTTCGGTAGGTGGAAGC 282
      |||||||

RESULT 16
US-10-357-930-4650/c
; Sequence 4650, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4650
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 106, 367, 418
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-4650

Query Match          6.7%; Score 85; DB 18; Length 438;
Best Local Similarity 100.0%; Pred. No. 8e-34;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 CTGGCCAGTGTTTAAAACTTCAGTTTTCACAGAAATAAGGACCCCATCTCTGTCGCCAAC 1249
      |||||||
Db 366 CTGGCCAGTGTTTAAAACTTCAGTTTTCACAGAAATAAGGACCCCATCTCTGTCGCCAAC 307
      |||||||

QY 1250 TAAAACTCTTCGGTAGGTGGAAGC 1274
      |||||||
Db 306 TAAAACTCTTCGGTAGGTGGAAGC 282
      |||||||

RESULT 17
US-09-960-352-10011
; Sequence 10011, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
```





```
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 612660
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-719-900-612660

Query Match          2.0%; Score 25; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1141 GGCATCAATGTAACATAATTTCCATGT 1165
      |||||
Db 1 GGCATCAATGTAACATAATTTCCATGT 25

RESULT 22
US-10-490-581-10/c
; Sequence 10, Application US/10490581
; Publication No. US20050070015A1
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Furukawa, Yoichi
; APPLICANT: Japan as represented by
; APPLICANT: The President of the University of Tokyo
; APPLICANT: Oncotherapy Science, Inc.
; TITLE OF INVENTION: Methods of Regulating Growth and Death of Cancer Cells
; FILE REFERENCE: 082368-000400US
; CURRENT APPLICATION NUMBER: US/10/490,581
; CURRENT FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: JP 2001-290248
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: WO PCT/JP02/08416
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: artificially
; OTHER INFORMATION: synthesized sequence
US-10-490-581-10

Query Match          1.9%; Score 24; DB 19; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1142 GGCATCAATGTAACATAATTTCCATGT 1165
      |||||
Db 33 GGCATCAATGTAACATAATTTCCATGT 10

RESULT 23
US-09-728-445-159
; Sequence 159, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020102543A1e1 Mutated Mammalian Cells and
; Animals

; FILE REFERENCE: LEX-0102-USA
; CURRENT APPLICATION NUMBER: US/09/728,445
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159
; LENGTH: 226
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(226)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-445-159

Query Match          1.8%; Score 23; DB 9; Length 226;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 AAGTGTCACCATCCAGAGGGT 386
      |||||
Db 175 AAGTGTCACCATCCAGAGGGT 197

RESULT 24
US-10-424-599-69723
; Sequence 69723, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 69723
; LENGTH: 2466
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33973C.1
US-10-424-599-69723

Query Match          1.7%; Score 22; DB 17; Length 2466;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 836 TTCCTGGTCTGTTGACTGGT 857
      |||||
Db 596 TTCCTGGTCTGTTGACTGGT 617

RESULT 25
US-10-490-581-5
; Sequence 5, Application US/10490581
; Publication No. US20050070015A1
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Furukawa, Yoichi
; APPLICANT: Japan as represented by
; APPLICANT: The President of the University of Tokyo
; APPLICANT: Oncotherapy Science, Inc.
; TITLE OF INVENTION: Methods of Regulating Growth and Death of Cancer Cells
; FILE REFERENCE: 082368-000400US
; CURRENT APPLICATION NUMBER: US/10/490,581
; CURRENT FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: JP 2001-290248
```

```
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: WO PCT/JP02/08416
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: artificially
; OTHER INFORMATION: synthesized sequence
US-10-490-581-5

Query Match          1.6%; Score 21; DB 19; Length 21;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 412 TCACCAATGACTTGGCGAGT 432
Db 1 TCACCAATGACTTGGCGAGT 21

RESULT 26
US-10-490-581-6/c
; Sequence 6, Application US/10490581
; Publication No. US20050070015A1
; GENERAL INFORMATION:
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Furuoka, Yoichi
; APPLICANT: Japan as represented by
; APPLICANT: The President of the University of Tokyo
; APPLICANT: Oncotherapy Science, Inc.
; TITLE OF INVENTION: Methods of Regulating Growth and Death of Cancer Cells
; FILE REFERENCE: 082368-00040005
; CURRENT APPLICATION NUMBER: US/10/490,581
; CURRENT FILING DATE: 2004-03-18
; PRIOR APPLICATION NUMBER: JP 2001-290248
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: WO PCT/JP02/08416
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: artificially
; OTHER INFORMATION: synthesized sequence
US-10-490-581-6

Query Match          1.6%; Score 21; DB 19; Length 21;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 644 CTCTGCCACACACAGAAAAG 664
Db 21 CTCTGCCACACACAGAAAAG 1

RESULT 27
US-10-719-900-37747
; Sequence 37747, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20

; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 257078
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-257078

Query Match          1.6%; Score 21; DB 19; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 859 ATGATGCAGTCTCTGTTTGGC 879
Db 1 ATGATGCAGTCTCTGTTTGGC 21

RESULT 29
US-09-918-995-15609
; Sequence 15609, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15609
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(442)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-15609

Query Match          1.6%; Score 21; DB 10; Length 442;
Best Local Similarity 100.0%; Pred. No. 6.4;
```

```
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 165 TTTTAAATATCTTTTAAAT 185
Db 307 TTTTAAATATCTTTTAAAT 327

RESULT 30
US-10-425-115-119327
; Sequence 119327, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASGO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: Q76319
; CURRENT APPLICATION NUMBER: US/10/627,757
; PRIOR FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 1094
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-10

Query Match 1.6%; Score 21; DB 17; Length 1094;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 165 TTTTAAATATCTTTTAAAT 185
Db 478 TTTTAAATATCTTTTAAAT 498

RESULT 31
US-10-627-757-9
; Sequence 9, Application US/10627757
; Publication No. US20040091914A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASGO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: Q76319
; CURRENT APPLICATION NUMBER: US/10/627,757
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-9

Query Match 1.6%; Score 21; DB 17; Length 1150;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 165 TTTTAAATATCTTTTAAAT 185
Db 1006 TTTTAAATATCTTTTAAAT 1026

RESULT 34
US-10-091-281-2
; Sequence 2, Application US/10091281
; Publication No. US20030190617A1
; GENERAL INFORMATION:
; APPLICANT: RAYMOND, VINCENT
; APPLICANT: SI, ERWIN
```

```
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 165 TTTTAAATATCTTTTAAAT 185
Db 307 TTTTAAATATCTTTTAAAT 327

RESULT 30
US-10-425-115-119327
; Sequence 119327, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASGO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: Q76319
; CURRENT APPLICATION NUMBER: US/10/627,757
; PRIOR FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 1094
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-10

Query Match 1.6%; Score 21; DB 17; Length 1094;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 165 TTTTAAATATCTTTTAAAT 185
Db 478 TTTTAAATATCTTTTAAAT 498

RESULT 31
US-10-627-757-9
; Sequence 9, Application US/10627757
; Publication No. US20040091914A1
; GENERAL INFORMATION:
; APPLICANT: KOUCHI YASUHIRO
; APPLICANT: MASASGO AKINORI
; APPLICANT: TAKAHATI TAKAYUKI
; TITLE OF INVENTION: GENE ASSAY METHOD FOR PREDICTING GLAUCOMA ONSET RISK
; FILE REFERENCE: Q76319
; CURRENT APPLICATION NUMBER: US/10/627,757
; CURRENT FILING DATE: 2003-07-28
; PRIOR APPLICATION NUMBER: JP P2002-226612
; PRIOR FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-627-757-9

Query Match 1.6%; Score 21; DB 17; Length 1150;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 165 TTTTAAATATCTTTTAAAT 185
Db 1006 TTTTAAATATCTTTTAAAT 1026

RESULT 34
US-10-091-281-2
; Sequence 2, Application US/10091281
; Publication No. US20030190617A1
; GENERAL INFORMATION:
; APPLICANT: RAYMOND, VINCENT
; APPLICANT: SI, ERWIN
```

```
; APPLICANT: MORISETTE, JEAN
; TITLE OF INVENTION: OPTINEURIN NUCLEIC ACID MOLECULES AND USES THEREOF
; FILE REFERENCE: 13587.338
; CURRENT APPLICATION NUMBER: US/10/091,281
; CURRENT FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 46951
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 391
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 691
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 709
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 887
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 834
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 987
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 1112
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 1505
; OTHER INFORMATION: insertion of additional c residue
; NAME/KEY: allele
; LOCATION: 1606
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 2405
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 2606
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 3113
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 3555
; OTHER INFORMATION: insertion of additional t residue
; NAME/KEY: allele
; LOCATION: 3625
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 3629
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 3882
; OTHER INFORMATION: insertion of additional t residue
; NAME/KEY: allele
; LOCATION: 3988
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: allele
; LOCATION: 4452
; OTHER INFORMATION: single nucleotide polymorphism (SNP)
; NAME/KEY: repeat_region
; LOCATION: 598..878
; OTHER INFORMATION: repeat element
; NAME/KEY: repeat_region
; LOCATION: 938..957
; OTHER INFORMATION: Short repeat element
; NAME/KEY: repeat_region
```

```
; LOCATION: 1002..1329
; OTHER INFORMATION: ALU repeat element
; NAME/KEY: repeat_region
; LOCATION: 2288..2587
; OTHER INFORMATION: ALU repeat element
; NAME/KEY: misc_feature
; LOCATION: 5054
; OTHER INFORMATION: putative transcription start site
US-10-091-281-2
Query Match 1.6%; Score 21; DB 16; Length 46951;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 165 TTTTAAATATCTTTTAAAT 185
|||||
Db 30767 TTTTAAATATCTTTTAAAT 30787
|||||
RESULT 35
US-10-679-246-42
; Sequence 42, Application US/10679246
; Publication No. US20040163138A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved
; TITLE OF INVENTION: in Protein Degradation, Products and Methods Related Thereto
; FILE REFERENCE: 6821-235
; CURRENT APPLICATION NUMBER: US/10/679,246
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US 09/591,694
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 28
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-10-679-246-42
Query Match 1.6%; Score 20; DB 18; Length 28;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 236 GGACTTATGGCATGTAACA 255
|||||
Db 9 GGACTTATGGCATGTAACA 28
|||||
RESULT 36
US-10-305-720-397
; Sequence 397, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 397
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 2113436
```

US-10-305-720-397

Query Match 1.6%; Score 20; DB 17; Length 257;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 850 GACTGGGTGATGATGCAGTC 869  
|||||  
Db 222 GACTGGGTGATGATGCAGTC 241

RESULT 37

US-10-242-535A-55073  
; Sequence 55073, Application US/10242535A  
; Publication No. US20040013663A1

; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; APPLICANT: Liew, C.C.

; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis

; FILE REFERENCE: 4231/2005

; CURRENT APPLICATION NUMBER: US/10/242,535A

; CURRENT FILING DATE: 2002-09-12

; PRIOR APPLICATION NUMBER: US 10/085,783

; PRIOR FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/275,017

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: US 60/271,955

; PRIOR FILING DATE: 2001-02-28

; NUMBER OF SEQ ID NOS: 58994

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 55073

; LENGTH: 335

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (6)..(6)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (9)..(9)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (20)..(20)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (65)..(65)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (318)..(319)

; OTHER INFORMATION: n is a, c, g, or t

US-10-242-535A-55073

Query Match

Best Local Similarity 1.6%; Score 20; DB 17; Length 335;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GTAATTTTATTTTAAATA 174  
|||||  
Db 43 GTAATTTTATTTTAAATA 62

RESULT 38

US-10-085-783A-55073

; Sequence 55073, Application US/10085783A

; Publication No. US20040037841A1

; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 55073

; LENGTH: 335

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (6)..(6)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (9)..(9)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (20)..(20)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (65)..(65)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (318)..(319)

; OTHER INFORMATION: n is a, c, g, or t

US-10-085-783A-55073

Query Match

Best Local Similarity 1.6%; Score 20; DB 17; Length 335;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 GTAATTTTATTTTAAATA 174  
|||||  
Db 43 GTAATTTTATTTTAAATA 62

RESULT 39

US-10-027-632-317039/c

; Sequence 317039, Application US/10027632

; Publication No. US20020198371A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

```

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 317039
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-317039

Query Match
Best Local Similarity 100.0%; Score 20; DB 13; Length 497;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TAAAGGACTTATGGCATGT 250
Db 159 TAAAGGACTTATGGCATGT 140

RESULT 40
US-10-027-632-317040/c
; Sequence 317040, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/195,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 317040
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-317040

Query Match
Best Local Similarity 100.0%; Score 20; DB 13; Length 497;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TAAAGGACTTATGGCATGT 250
Db 159 TAAAGGACTTATGGCATGT 140

RESULT 41
US-10-027-632-317039/c
; Sequence 317039, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```

; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 317039
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-317039

Query Match
Best Local Similarity 100.0%; Score 20; DB 17; Length 497;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TAAAGGACTTATGGCATGT 250
Db 159 TAAAGGACTTATGGCATGT 140

RESULT 42
US-10-027-632-317040/c
; Sequence 317040, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 317040
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-317040

Query Match
Best Local Similarity 100.0%; Score 20; DB 17; Length 497;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TAAAGGACTTATGGCATGT 250
Db 159 TAAAGGACTTATGGCATGT 140

RESULT 43
US-10-027-632-88621/c
; Sequence 88621, Application US/10027632
```

; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 88621  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-88621

Query Match 1.6%; Score 20; DB 13; Length 498;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TAAAGGACTTATGGCATGT 250  
Db 159 TAAAGGACTTATGGCATGT 140

RESULT 44  
US-10-027-632-88622/c  
; Sequence 88622, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 88622  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-88622

Query Match 1.6%; Score 20; DB 13; Length 498;

Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 231 TAAAGGACTTATGGCATGT 250  
Db 159 TAAAGGACTTATGGCATGT 140

RESULT 45  
US-10-027-632-88621/c  
; Sequence 88621, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 88621  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-88621

Query Match 1.6%; Score 20; DB 17; Length 498;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TAAAGGACTTATGGCATGT 250  
Db 159 TAAAGGACTTATGGCATGT 140

RESULT 46  
US-10-027-632-88622/c  
; Sequence 88622, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28



; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 88622  
; LENGTH: 498  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-88622

Query Match 1.6%; Score 20; DB 17; Length 498;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 TAAAGGACTTATGGCATCT 250  
DB 159 TAAAGGACTTATGGCATCT 140

RESULT 47  
US-10-027-632-91273/c  
; Sequence 91273, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 91273  
; LENGTH: 507  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-91273

Query Match 1.6%; Score 20; DB 13; Length 507;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TGTAAATTTATTTTAAAT 173  
DB 407 TGTAAATTTATTTTAAAT 388

RESULT 48  
US-10-027-632-317676/c  
; Sequence 317676, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 317676  
; LENGTH: 507  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-317676

Query Match 1.6%; Score 20; DB 13; Length 507;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TGTAAATTTATTTTAAAT 173  
DB 407 TGTAAATTTATTTTAAAT 388

RESULT 49  
US-10-027-632-91273/c  
; Sequence 91273, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 91273  
; LENGTH: 507  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-91273

Query Match 1.6%; Score 20; DB 17; Length 507;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TGTAAATTTATTTTAAAT 173  
DB 407 TGTAAATTTATTTTAAAT 388

RESULT 50

US-10-027-632-317676/c  
; Sequence 317676, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 317676  
; LENGTH: 507  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-317676

Query Match 1.6%; Score 20; DB 17; Length 507;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 154 TGTAAATTTTATTTTAAAT 173  
|||  
Db 407 TGTAAATTTTATTTTAAAT 388

Search completed: April 25, 2005, 06:36:57  
Job time : 2403 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 03:56:57 ; Search time 80 Seconds  
(without alignments)  
1440.682 Million cell updates/sec

Title: US-10-679-246-2

Perfect score: 1611

Sequence: 1 MWIIIFLLPPYVFISEMSRQ.....IAQLFAENGLGIVTISMC 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	298	4 AAB35157	Aab35157 Human Sia
2	1603	99.5	298	7 ADM04037	Adm04037 Human pro
3	1537	95.4	313	8 ADS34491	Ads34491 POSH prot
4	1532	95.1	282	8 ADS34492	Ads34492 POSH prot
5	1528	94.8	282	5 ABB84454	Abb84454 Murine mm
6	1528	94.8	282	5 ABB84454	Abb84454 Mouse isc
7	1522	94.5	282	5 ABB84452	Abb84452 Human hss
8	1498	93.0	282	5 ABB84455	Abb84455 Murine mm
9	1489	92.4	282	2 AAW18520	Aaw18520 Tumour su
10	1198	74.4	324	5 ABB84447	Abb84447 Human Sai
11	1198	74.4	324	5 ABB84453	Abb84453 Human hss
12	1198	74.4	324	6 ABR47593	Abr47593 Breast ca
13	1190	73.9	324	8 ADR14123	Adr14123 Human NF-
14	1190	73.9	324	8 ADP54831	Adp54831 Human PRO
15	1190	73.9	325	5 ABB84456	Abb84456 Murine mm
16	1190	73.9	325	5 ABB57147	Abb57147 Mouse isc
17	1153.5	71.6	314	4 ABB84778	Abb84778 Drosophi
18	1153.5	71.6	314	5 ABB84451	Abb84451 Siah-rela
19	1037.5	64.4	339	5 ABB84450	Abb84450 Siah-rela
20	530	32.9	310	4 ABB67709	Abb67709 Drosophi
21	447	27.7	284	3 AAG33121	Aag33121 Zea mays
22	447	27.7	313	3 AAV67582	Aav67582 Maize SIN
23	447	27.7	313	3 AAG33120	Aag33120 Zea mays
24	439	27.3	315	5 ABB84458	Abb84458 Siah-rela
25	434.5	27.0	297	3 AAG20022	Aag20022 Arabidops

26	434.5	27.0	297	3	AAG54067	Aag54067 Arabidops
27	434.5	27.0	308	3	AAG54066	Aag54066 Arabidops
28	434.5	27.0	308	3	AAG20021	Aag20021 Arabidops
29	432.5	26.8	290	3	AAG20023	Aag20023 Arabidops
30	432.5	26.8	290	3	AAG54068	Aag54068 Arabidops
31	432	26.8	89	3	AAB54091	Aab54091 Human pan
32	429	26.6	322	3	AAG31375	Aag31375 Arabidops
33	429	26.6	327	3	AAG31374	Aag31374 Arabidops
34	429	26.6	327	5	ABB84449	Abb84449 Siah-rela
35	423	26.3	321	3	AAG13847	Aag13847 Arabidops
36	423	26.3	326	3	AAG13846	Aag13846 Arabidops
37	423	26.3	361	3	AAG13845	Aag13845 Arabidops
38	422	26.2	305	3	AAG30769	Aag30769 Arabidops
39	422	26.2	305	5	ABB84448	Abb84448 Siah-rela
40	421	26.1	281	3	AAG30770	Aag30770 Arabidops
41	421	26.1	349	3	AAG35743	Aag35743 Zea mays
42	421	26.1	349	3	AAG35742	Aag35742 Zea mays
43	421	26.1	373	3	AAG35741	Aag35741 Zea mays
44	419	26.0	321	3	AAG53803	Aag53803 Arabidops
45	419	26.0	326	3	AAG53802	Aag53802 Arabidops

ALIGNMENTS

RESULT 1  
AAB35157  
ID AAB35157 standard; protein; 298 AA.  
XX  
AC AAB35157;  
XX  
DT 09-APR-2001 (first entry)  
XX  
DE Human Siah-1alpha SEQ ID NO: 2.  
XX  
KW Human; protein degradation; siah-mediated degradation protein; SMDF;  
KW SCF-complex protein; SCP; siah-1alpha; siah-1 interacting protein; SIP;  
KW Skp1-associated F-box protein; SAF-1; SAF-2; SAD; cancer; cell division;  
KW Skp1-associated destruction-box protein; inflammatory disease.  
XX  
OS Homo sapiens.  
XX  
FN WO200077207-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 09-JUN-2000; 2000WO-US015873.  
XX  
PR 11-JUN-1999; 99US-00330517.  
XX  
PA (BURN-) BURNHAM INST.  
XX  
PI Reed JC, Matsuzawa S;  
XX  
WP1; 2001-071273/08.  
DR N-PSDB; AAG67281.  
XX  
PT Siah-Mediated Degradation Protein, useful for drug screening, for  
XX therapeutic applications and for functional genomics.  
XX  
PS Claim 15; Page 97; 121pp; English.  
XX  
CC The present invention provides the protein and coding sequences of  
CC several siah-mediated degradation proteins and SCF-complex proteins.  
CC These are designated Siah-1alpha, Siah-1 interacting protein (SIP), which  
CC encodes two proteins due to alternative splicing (SIP-L and SIP-S), Skp1-  
CC associated F-box protein-lalpha and beta and -2 (SAF-lalpha, SAF-lbeta  
CC and SAF-2) and Skp1-associated destruction-box protein (SAD). The  
CC proteins and their coding sequences are useful in the diagnosis and  
CC treatment of cancers, disorders where too little cell division occurs  
CC such as bone marrow aplasia, immunodeficiencies and inflammatory  
CC diseases including sepsis, fibrosis, arthritis and graft versus host  
CC disease

CC are useful as pharmaceutical agents. The present sequence represents a  
CC protein sequence of the invention.  
XX  
SQ Sequence 298 AA;  
Query Match 99.5%; Score 1603; DB 7; Length 298;  
Best Local Similarity 99.7%; Pred. No. 4.8e-149;  
Matches 297; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MVIIFLLPPVVFISEMSRQTATALTPTGTSKPPSQRPVLPALTTGTASNNDLASLFECPVC 60  
DB 1 MVIIFLLPPVVFISEMSRQTATALTPTGTSKPPSQRPVLPALTTGTASNNDLASLFECPVC 60  
QY 61 FDYVLPPILOQSGHLVCSNCRPKLTCCPTCRGLGSRINLAMEKVANSVLPCKYASSG 120  
DB 61 FDYVLPPILOQSGHLVCSNCRPKLTCCPTCRGLGSRINLAMEKVANSVLPCKYASSG 120  
QY 121 CEITLPHTEKADHEELCEFRPYSCPCGASCKWQSGSLDAMPHLMHQHKSITTLQGEDIV 180  
DB 121 CEITLPHTEKADHEELCEFRPYSCPCGASCKWQSGSLDAMPHLMHQHKSITTLQGEDIV 180  
QY 181 FLATDINLPGAVDWMVQSCFGFHFMLVLEKQKYDGHQOQFFAIVQLIGTRKQAFAYR 240  
DB 181 FLATDINLPGAVDWMVQSCFGFHFMLVLEKQKYDGHQOQFFAIVQLIGTRKQAFAYR 240  
QY 241 LELNGHRRRLTWEATPRSIHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISM 298  
DB 241 LELNGHRRRLTWEATPRSIHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISM 298  
RESULT 3  
ADS34491  
ID ADS34491 standard; protein; 313 AA.  
XX  
AC ADS34491;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE POSH protein associated protein #55.  
XX  
KW cytostatic; nootropic; neuroprotective; antiparkinsonian; anticonvulsant;  
KW antiviral; neuroleptic; central nervous system; POSH polypeptide;  
KW POSH-associated protein; POSH-AP; HERPUNDI; Ubiquitin ligase;  
KW antiviral agent; anti-apoptotic agent; anti-cancer agent;  
KW secretory pathway trafficking inhibitor;  
KW neurological disorder progression disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; schizophrenia;  
KW Niemann-Pick's disease.  
XX  
OS Homo sapiens.  
XX  
PN WO2004078130-A2.  
XX  
PD 16-SEP-2004.  
XX  
PF 02-MAR-2004; 2004WO-US006308.  
XX  
PR 03-MAR-2003; 2003US-0451437P.  
PR 05-MAR-2003; 2003US-0452284P.  
PR 19-MAR-2003; 2003US-0455760P.  
PR 20-MAR-2003; 2003US-0456640P.  
PR 03-APR-2003; 2003US-0460526P.  
PR 04-APR-2003; 2003US-0460792P.  
PR 21-APR-2003; 2003US-0464285P.  
PR 09-MAY-2003; 2003US-0469462P.  
PR 15-MAY-2003; 2003US-0471378P.  
PR 20-MAY-2003; 2003US-0472327P.  
PR 30-MAY-2003; 2003US-0474706P.  
PR 03-JUN-2003; 2003US-0475825P.  
PR 17-JUN-2003; 2003US-0479317P.  
PR 19-JUN-2003; 2003US-0480215P.  
PR 19-JUN-2003; 2003US-0480376P.  
PR 08-AUG-2003; 2003US-0493860P.

XX  
SQ Sequence 298 AA;  
Query Match 100.0%; Score 1611; DB 4; Length 298;  
Best Local Similarity 100.0%; Pred. No. 7.8e-150;  
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVIIFLLPPVVFISEMSRQTATALTPTGTSKPPSQRPVLPALTTGTASNNDLASLFECPVC 60  
DB 1 MVIIFLLPPVVFISEMSRQTATALTPTGTSKPPSQRPVLPALTTGTASNNDLASLFECPVC 60  
QY 61 FDYVLPPILOQSGHLVCSNCRPKLTCCPTCRGLGSRINLAMEKVANSVLPCKYASSG 120  
DB 61 FDYVLPPILOQSGHLVCSNCRPKLTCCPTCRGLGSRINLAMEKVANSVLPCKYASSG 120  
QY 121 CEITLPHTEKADHEELCEFRPYSCPCGASCKWQSGSLDAMPHLMHQHKSITTLQGEDIV 180  
DB 121 CEITLPHTEKADHEELCEFRPYSCPCGASCKWQSGSLDAMPHLMHQHKSITTLQGEDIV 180  
QY 181 FLATDINLPGAVDWMVQSCFGFHFMLVLEKQKYDGHQOQFFAIVQLIGTRKQAFAYR 240  
DB 181 FLATDINLPGAVDWMVQSCFGFHFMLVLEKQKYDGHQOQFFAIVQLIGTRKQAFAYR 240  
QY 241 LELNGHRRRLTWEATPRSIHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISM 298  
DB 241 LELNGHRRRLTWEATPRSIHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISM 298  
RESULT 2  
ADM04027  
ID ADM04027 standard; protein; 298 AA.  
XX  
AC ADM04027;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Human protein of the invention SEQ ID NO:2712.  
XX  
KW human; gene therapy; diagnostic marker; pharmaceutical.  
XX  
OS Homo sapiens.  
XX  
PN EP1347046-A1.  
XX  
PD 24-SEP-2003.  
XX  
PF 12-APR-2002; 2002EP-00008400.  
XX  
PR 22-MAR-2002; 2002JP-00137785.  
XX  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
DR WPI; 2003-723558/69.  
DR N-PSDB; ADM01584.  
XX  
PT New polynucleotides and polypeptides are useful in gene therapy, for  
PT developing a diagnostic marker or medicines for regulating their  
PT expression and activity, or as a target of gene therapy.  
XX  
PS Claim 1; SEQ ID NO 2712; 305pp; English.  
XX  
CC The invention relates to a novel human polynucleotide and the encoded  
CC polypeptide. A polynucleotide of the invention may have a use in gene  
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful  
CC as a primer for synthesizing the polynucleotide or as a probe for  
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are  
CC useful in gene therapy, for developing a diagnostic marker or medicines  
CC for regulating their expression and activity, or as a target of gene  
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides

PR 28-AUG-2003; 2003US-0498634P.  
 PR 16-SEP-2003; 2003US-0503931P.  
 PR 10-NOV-2003; 2003WO-US035712.  
 PR 05-FEB-2004; 2004WO-US003600.  
 PR 02-MAR-2004; 2004US-0549896P.  
 XX  
 PA (PROT-) PROTEOLOGICS INC.  
 XX  
 XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
 PI Greener T;  
 XX WPI; 2004-662346/64.  
 XX  
 XX Isolated, purified or recombinant complex, useful for identifying an  
 PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
 PT POSH-associated protein (POSH-AP).  
 XX  
 XX Disclosure; SEQ ID NO 255; 374pp; English.  
 PS  
 XX The invention relates to an isolated, purified or recombinant complex (I)  
 CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
 CC or HERPUDI and a ubiquitin ligase (b). Methods using (I), (a) or (b) are  
 CC useful for identifying an agent that modulates an activity of a POSH  
 CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
 CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
 CC of a protein through the secretory pathway, an agent that inhibits the  
 CC progression of a neurological disorder, an agent that modulates a POSH  
 CC function, an agent that modulates a HERPUDI function. The methods can be  
 CC used for treating a viral infection, for inhibiting an activity of a POSH  
 CC -AP in a cell, for treating a POSH-associated disease in a subject. The  
 CC POSH-associated disease is viral infection, POSH-associated cancer or  
 CC POSH-associated neurological disorder. The methods are useful for  
 CC treating or preventing POSH-associated neurological disorder in a subject  
 CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
 CC protein of the invention.  
 XX  
 XX Sequence 313 AA;  
 SQ  
 Query Match 95.4%; Score 1537; DB 8; Length 313;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-142;  
 Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 16 EMSQTATALPTGTSKCPSPQRPVLTGTTASNNDLASLFCPCFVDYVLPPILOCSGH 75  
 DB 31 EMSQTATALPTGTSKCPSPQRPVLTGTTASNNDLASLFCPCFVDYVLPPILOCSGH 90  
 QY 76 LVCSNCRPKLTCCPTCRGLSIRNLAMEKVANSVLPCKYASSGCEITLPTKADHDEE 135  
 DB 91 LVCSNCRPKLTCCPTCRGLSIRNLAMEKVANSVLPCKYASSGCEITLPTKADHDEE 150  
 QY 136 LCEFRPSPCPGACCKWQSLDVMPLMHQHSITTLQGEDIVFLATDINLPADVWV 195  
 DB 151 LCEFRPSPCPGACCKWQSLDVMPLMHQHSITTLQGEDIVFLATDINLPADVWV 210  
 QY 196 MMQSCFGFHMVLVEKQEKYDGHQQFAIVOLIGTRKQAFNAYRLLENGHRRRLTWEAT 255  
 DB 211 MMQSCFGFHMVLVEKQEKYDGHQQFAIVOLIGTRKQAFNAYRLLENGHRRRLTWEAT 270  
 QY 256 PRSIEGIIATAMNSDCLVFDTSIAQLFNGNLGINVTISM 298  
 DB 271 PRSIEGIIATAMNSDCLVFDTSIAQLFNGNLGINVTISM 313  
 RESULT 4  
 ADS34492  
 ID ADS34492 standard; protein; 282 AA.  
 XX  
 AC ADS34492;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE POSH protein associated protein #56.

XX cytostatic; neurotropic; neuroprotective; antiparkinsonian; anticonvulsant;  
 KW antiviral; neuroleptic; central nervous system; POSH polypeptide;  
 KW POSH-associated protein; POSH-AP; HERPUDI; Ubiquitin ligase;  
 KW antiviral agent; anti-apoptotic agent; anti-cancer agent;  
 KW secretory pathway trafficking inhibitor;  
 KW neurological disorder progression disorder; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease; schizophrenia;  
 KW Niemann-Pick's disease.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2004078130-A2.  
 XX  
 XX 16-SEP-2004.  
 XX  
 XX 02-MAR-2004; 2004WO-US006308.  
 XX  
 XX 03-MAR-2003; 2003US-0451437P.  
 PR 05-MAR-2003; 2003US-0452284P.  
 PR 19-MAR-2003; 2003US-0455760P.  
 PR 20-MAR-2003; 2003US-0456640P.  
 PR 03-APR-2003; 2003US-0460526P.  
 PR 04-APR-2003; 2003US-0460792P.  
 PR 21-APR-2003; 2003US-0464285P.  
 PR 09-MAY-2003; 2003US-0469462P.  
 PR 15-MAY-2003; 2003US-0471378P.  
 PR 20-MAY-2003; 2003US-0472327P.  
 PR 30-MAY-2003; 2003US-0474706P.  
 PR 03-JUN-2003; 2003US-0475825P.  
 PR 17-JUN-2003; 2003US-0479317P.  
 PR 19-JUN-2003; 2003US-0480215P.  
 PR 08-AUG-2003; 2003US-0493860P.  
 PR 28-AUG-2003; 2003US-0498634P.  
 PR 16-SEP-2003; 2003US-0503931P.  
 PR 10-NOV-2003; 2003WO-US035712.  
 PR 05-FEB-2004; 2004WO-US003600.  
 PR 02-MAR-2004; 2004US-0549896P.  
 XX  
 XX (PROT-) PROTEOLOGICS INC.  
 XX  
 XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
 PI Greener T;  
 XX WPI; 2004-662346/64.  
 XX  
 XX Isolated, purified or recombinant complex, useful for identifying an  
 PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and  
 PT POSH-associated protein (POSH-AP).  
 XX  
 XX Disclosure; SEQ ID NO 256; 374pp; English.  
 PS  
 XX The invention relates to an isolated, purified or recombinant complex (I)  
 CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)  
 CC or HERPUDI and a ubiquitin ligase (b). Methods using (I), (a) or (b) are  
 CC useful for identifying an agent that modulates an activity of a POSH  
 CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-  
 CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking  
 CC of a protein through the secretory pathway, an agent that inhibits the  
 CC progression of a neurological disorder, an agent that modulates a POSH  
 CC function, an agent that modulates a HERPUDI function. The methods can be  
 CC used for treating a viral infection, for inhibiting an activity of a POSH  
 CC -AP in a cell, for treating a POSH-associated disease in a subject. The  
 CC POSH-associated disease is viral infection, POSH-associated cancer or  
 CC POSH-associated neurological disorder. The methods are useful for  
 CC treating or preventing POSH-associated neurological disorder in a subject  
 CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,  
 CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a  
 CC protein of the invention.  
 XX  
 XX Sequence 282 AA;  
 SQ

Query Match 95.1%; Score 1532; DB 8; Length 282;  
Best Local Similarity 100.0%; Pred. No. 4.4e-142;  
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 MSRTATATPTGTSKCPSPQSRVPALTGTASNNDLASLFECPCVDFDYVLPPILOQSGHL 76  
DB 1 MSRTATATPTGTSKCPSPQSRVPALTGTASNNDLASLFECPCVDFDYVLPPILOQSGHL 60

QY 77 VCSNCRPKLTCCPTCRGPIGSRINLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136  
DB 61 VCSNCRPKLTCCPTCRGPIGSRINLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120

QY 137 CEFRPYSCPCPGASCKWQSLDAMVPHLMHOKHSITTLQGEDIVFLATDINLPGAVDWM 196  
DB 121 CEFRPYSCPCPGASCKWQSLDAMVPHLMHOKHSITTLQGEDIVFLATDINLPGAVDWM 180

QY 197 MQSCFGFHFMLVLEKQEKYDGHQQFFAIVQLIGTRKQAFNFAIRLELNGHRRRLTWEATP 256  
DB 181 MQSCFGFHFMLVLEKQEKYDGHQQFFAIVQLIGTRKQAFNFAIRLELNGHRRRLTWEATP 240

QY 257 RSIHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISMC 298  
DB 241 RSIHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISMC 282

RESULT 5  
ABB84454  
ID ABB84454 standard; peptide; 282 AA.  
AC ABB84454;  
XX  
DT 08-NOV-2002 (first entry)  
XX  
DE Murine mmSiah1A protein.  
XX  
KW mmSiah1A; seven in absentia homologue; ubiquitination pathway; cancer;  
KW infertility; inflammation; Siah; cytostatic; antiinflammatory;  
KW neuroprotective; immunosuppressive; apoptosis; NF kappa B signaling;  
KW neurological disorder; protein co-ordinate data.  
XX  
OS Mus musculus.  
XX  
XX WO200262838-A1.  
XX  
XX 15-AUG-2002.  
XX  
XX 06-FEB-2002; 2002WO-AU000118.  
XX  
XX 06-FEB-2001; 2001AU-00002908.  
XX  
XX (MACC-) MACCALLUM CANCER INST PETER.  
XX (SVIN-) ST VINCENTS INST MEDICAL RES.  
XX  
XX Bowtell DDL, House CM, Parker MW, Polekhina G;  
XX PI  
XX WPI; 2002-657521/70.  
XX  
XX  
XX New binding domain of a seven in absentia homolog protein, useful for  
XX identifying agonists or antagonists of the domain, for treating or  
XX preventing e.g. cancer, inflammation, infertility or other related  
XX conditions, or apoptosis.  
XX  
XX Disclosure; Fig 2; 153pp; English.  
XX  
XX This invention describes a novel binding domain or a portion of the  
XX binding domain of a seven in absentia homologue (Siah) protein. Siah is  
XX an important protein in the ubiquitination pathway and the binding domain  
XX is capable of binding substrates, co-factors and interactors of the Siah  
XX protein. The products of the invention have antiinflammatory, cytostatic,  
XX neuroprotective and immunosuppressive activity. The binding domain is  
XX useful for identifying agonists and antagonists to the domain. The  
XX antagonist, agonist or ligand of the binding domain of Siah may be used  
XX for treating or preventing cancer e.g. breast cancer, inflammation,

CC infertility or other related conditions, a disease relating to abnormal  
CC protein degradation, a pathological immune response, a disease relating  
CC to apoptosis, a disease relating to NF kappa B signaling, or a  
CC neurological disorder. This sequence represents a Siah-related protein,  
CC mmSiah1A described in the disclosure of the invention. Note: This  
CC sequence contains 3-D protein co-ordinate data  
XX  
XX Sequence 282 AA;

Query Match 94.8%; Score 1528; DB 5; Length 282;  
Best Local Similarity 99.6%; Pred. No. 1.1e-141;  
Matches 281; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 MSRTATATPTGTSKCPSPQSRVPALTGTASNNDLASLFECPCVDFDYVLPPILOQSGHL 76  
DB 1 MSRTATATPTGTSKCPSPQSRVPALTGTASNNDLASLFECPCVDFDYVLPPILOQSGHL 60

QY 77 VCSNCRPKLTCCPTCRGPIGSRINLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136  
DB 61 VCSNCRPKLTCCPTCRGPIGSRINLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120

QY 137 CEFRPYSCPCPGASCKWQSLDAMVPHLMHOKHSITTLQGEDIVFLATDINLPGAVDWM 196  
DB 121 CEFRPYSCPCPGASCKWQSLDAMVPHLMHOKHSITTLQGEDIVFLATDINLPGAVDWM 180

QY 197 MQSCFGFHFMLVLEKQEKYDGHQQFFAIVQLIGTRKQAFNFAIRLELNGHRRRLTWEATP 256  
DB 181 MQSCFGFHFMLVLEKQEKYDGHQQFFAIVQLIGTRKQAFNFAIRLELNGHRRRLTWEATP 240

QY 257 RSIHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISMC 298  
DB 241 RSIHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISMC 282

RESULT 6  
ABB57146  
ID ABB57146 standard; protein; 282 AA.  
AC ABB57146;  
XX  
XX 07-MAR-2002 (first entry)  
XX  
XX Mouse ischaemic condition related protein sequence SEQ ID NO:355.  
XX  
XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
XX vasospastic ischaemia; ischaemic condition; ischaemic disease.  
XX  
XX Mus musculus.  
XX  
XX WO200188188-A2.  
XX  
XX 22-NOV-2001.  
XX  
XX 18-MAY-2001; 2001WO-JP0004192.  
XX  
XX 18-MAY-2000; 2000JP-00145977.  
XX  
XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.  
XX  
XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;  
XX WPI; 2002-034733/04.  
XX  
XX N-PSDB; ABI99429.  
XX  
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
XX expression levels of particular genes defined in the specification or by  
XX determining the expression profile of a gene group comprising these  
XX genes.  
XX  
XX Claim 2; Page 958-959; 2690pp; English.  
XX  
XX The present invention describes a method for examining ischaemic  
XX conditions, comprising measuring the expression levels of particular



CC genes (I) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (I). The method  
 CC is useful for examining the ischaemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
 CC expression levels of particular genes (ABI99912 to ABI99914, encoding the  
 CC protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The expression  
 CC levels or expression profiles produced by these genes are used as an  
 CC indicator when screening for ischaemic condition-improving drugs or  
 CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR  
 CC primers for a mouse ischaemic condition related sequence, which are used  
 CC in the exemplification of the present invention  
 XX  
 SQ Sequence 282 AA;

Query Match 94.8%; Score 1528; DB 5; Length 282;  
 Best Local Similarity 99.6%; Pred. No. 1.1e-141;  
 Matches 281; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 17 MSRTATATPTGTSKCPSPQSRVPALGTGTASNNDLASLFECPVCFDYVLPPILOCSGHL 76  
 Db 1 MSRTATATPTGTSKCPSPQSRVPALGTGTASNNDLASLFECPVCFDYVLPPILOCSGHL 60  
 QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASGCEITLPHTEKADHEEL 136  
 Db 61 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASGCEITLPHTEKADHEEL 120  
 QY 137 CEFRPYSCPCGASCKWQSLDAMVPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 196  
 Db 121 CEFRPYSCPCGASCKWQSLDAMVPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 180  
 QY 197 MOSCFGHFMVLEKQEKYDGHQOFPAIVOLIGTRKQAEAFYRLNGLHRRRLTWEATP 256  
 Db 181 MOSCFGHFMVLEKQEKYDGHQOFPAIVOLIGTRKQAEAFYRLNGLHRRRLTWEATP 240  
 QY 257 RSIEGIATAIMNSDCLVFTSIAQLFAENGNGINVTISM 298  
 Db 241 RSIEGIATAIMNSDCLVFTSIAQLFAENGNGINVTISM 282

RESULT 7  
 ABB84452  
 ID ABB84452 standard; peptide; 282 AA.  
 XX  
 AC ABB84452;  
 XX  
 DT 08-NOV-2002 (first entry)  
 XX  
 DE Human hSiah1 protein.  
 XX  
 KW hSiah1; seven in absentia homologue; ubiquitination pathway; cancer;  
 XX infertility; inflammation; Siah; cytostatic; antiinflammatory;  
 KW neuroprotective; immunosuppressive; apoptosis; NF kappa B signaling;  
 KW neurological disorder; protein co-ordinate data.

XX Homo sapiens.  
 XX WO200262838-A1.  
 XX  
 PD 15-AUG-2002.  
 XX  
 XX 06-FEB-2002; 2002WO-AU000118.  
 XX  
 XX 06-FEB-2001; 2001AU-00002908.  
 XX  
 PA (MACC-) MACCALLUM CANCER INST PETER.  
 XX (SVIN-) ST VINCENTS INST MEDICAL RES.  
 XX  
 PI Bowtell DDL, House CM, Parker MW, Polekhina G;  
 XX WPI; 2002-657521/70.  
 DR  
 XX  
 XX New binding domain of a seven in absentia homolog protein, useful for

PT Identifying agonists or antagonists of the domain, for treating or  
 PT preventing e.g. cancer, inflammation, infertility or other related  
 PT conditions, or apoptosis.  
 XX  
 PS Disclosure; Fig 2; 153pp; English.  
 XX  
 CC This invention describes a novel binding domain or a portion of the  
 CC binding domain of a seven in absentia homologue (Siah) protein. Siah is  
 CC an important protein in the ubiquitination pathway and the binding domain  
 CC is capable of binding substrates, co-factors and interactors of the Siah  
 CC protein. The products of the invention have antiinflammatory, cytostatic,  
 CC neuroprotective and immunosuppressive activity. The binding domain is  
 CC useful for identifying agonists and antagonists to the domain. The  
 CC antagonist, agonist or ligand of the binding domain of Siah may be used  
 CC for treating or preventing cancer e.g. breast cancer, inflammation,  
 CC infertility or other related conditions, a disease relating to abnormal  
 CC protein degradation, a pathological immune response, a disease relating  
 CC to apoptosis, a disease relating to NF kappa B signaling, or a  
 CC neurological disorder. This sequence represents a Siah-related protein,  
 CC hSiah described in the disclosure of the invention. Note: This sequence  
 CC contains 3-D protein co-ordinate data  
 XX  
 SQ Sequence 282 AA;

Query Match 94.5%; Score 1522; DB 5; Length 282;  
 Best Local Similarity 98.9%; Pred. No. 4.2e-141;  
 Matches 279; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 17 MSRTATATPTGTSKCPSPQSRVPALGTGTASNNDLASLFECPVCFDYVLPPILOCSGHL 76  
 Db 1 MSRTATATPTGTSKCPSPQSRVPALGTGTASNNDLASLFECPVCFDYVLPPILOCSGHL 60  
 QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASGCEITLPHTEKADHEEL 136  
 Db 61 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASGCEITLPHTEKADHEEL 120  
 QY 137 CEFRPYSCPCGASCKWQSLDAMVPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 196  
 Db 121 CEFRPYSCPCGASCKWQSLDAMVPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 180  
 QY 197 MOSCFGHFMVLEKQEKYDGHQOFPAIVOLIGTRKQAEAFYRLNGLHRRRLTWEATP 256  
 Db 181 MOSCFGHFMVLEKQEKYDGHQOFPAIVOLIGTRKQAEAFYRLNGLHRRRLTWEATP 240  
 QY 257 RSIEGIATAIMNSDCLVFTSIAQLFAENGNGINVTISM 298  
 Db 241 RSIEGIATAIMNSDCLVFTSIAQLFAENGNGINVTISM 282

RESULT 8  
 ABB84455  
 ID ABB84455 standard; peptide; 282 AA.  
 XX  
 AC ABB84455;  
 XX  
 DT 08-NOV-2002 (first entry)  
 XX  
 DE Murine mmSiah1B protein.  
 XX  
 KW mmSiah1B; seven in absentia homologue; ubiquitination pathway; cancer;  
 XX infertility; inflammation; Siah; cytostatic; antiinflammatory;  
 KW neuroprotective; immunosuppressive; apoptosis; NF kappa B signaling;  
 KW neurological disorder; protein co-ordinate data.

XX Mus musculus.  
 XX WO200262838-A1.  
 XX  
 PD 15-AUG-2002.  
 XX  
 XX 06-FEB-2002; 2002WO-AU000118.  
 XX  
 XX 06-FEB-2001; 2001AU-00002908.

XX (MACC-) MACCALLUM CANCER INST PETER.  
XX (SVIN-) ST VINCENTS INST MEDICAL RES.  
XX  
XX Bowtell DDL, House CM, Parker MW, Polekhina G;  
XX WPI; 2002-657521/70.  
XX  
XX New binding domain of a seven in absentia homolog protein, useful for  
XX identifying agonists or antagonists of the domain, for treating or  
XX preventing e.g. cancer, inflammation, infertility or other related  
XX conditions, or apoptosis.  
XX  
XX Disclosure; Fig 2; i33pp; English.  
XX  
XX This invention describes a novel binding domain or a portion of the  
XX binding domain of a seven in absentia homologue (Siah) protein. Siah is  
XX an important protein in the ubiquitination pathway and the binding domain  
XX is capable of binding substrates, co-factors and interactors of the Siah  
XX protein. The products of the invention have antiinflammatory, cytostatic,  
XX neuroprotective and immunosuppressive activity. The binding domain is  
XX useful for identifying agonists and antagonists to the domain. The  
XX antagonist, agonist or ligand of the binding domain of Siah may be used  
XX for treating or preventing cancer e.g. breast cancer, inflammation,  
XX infertility or other related conditions, a disease relating to abnormal  
XX protein degradation, a pathological immune response, a disease relating  
XX to apoptosis, a disease relating to NF kappa B signaling, or a  
XX neurological disorder. This sequence represents a Siah-related protein,  
XX mmsiah1B described in the disclosure of the invention. Note: This  
XX sequence contains 3-D protein co-ordinate data  
XX  
XX Sequence 282 AA;  
SQ

Query Match 93.0%; Score 1498; DB 5; Length 282;  
Best Local Similarity 97.5%; Pred. No. 9.7e-139;  
Matches 275; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 17 MSRQTATALTGTSTKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQOSGHL 76  
DB 1 MSRQAATALTSTGTSTKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQOSGHL 60

QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136  
DB 61 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120

QY 137 CEFRPYSCPCPGASCCKWQSLDAMVPHLMHQHKSITTLQGEDIVFLATDINLPQAVDWM 196  
DB 121 CEFRPYSCPCPGASCCKWQSLDAMVPHLMHQHKSITTLQGEDIVFLATDINLPQAVDWM 180

QY 197 MQSCFGFHFMLVLEKQKYDGHQOFFAIVOLIIGTRKQAEAFYRLINGHRRRLTWEATP 256  
DB 181 MQSCFGFHFMLVLEKQKYDGHQOFFAIVOLIIGTRKQAEAFYRLINGHRRRLTWEATP 240

QY 257 RSIHEGIATAIMNSDCLVFTDTSIAQLFAENGNGINVTISM 298  
DB 241 RSIHEGIATAIMNSDCLVFTDTSIAQLFAENGNGINVTISM 282

RESULT 9  
AAW18520  
ID AAW18520 standard; protein; 282 AA.  
XX  
XX AAW18520;  
XX  
XX 25-MAR-1998 (first entry)  
XX Tumour suppressor activated pathway TSAP3 protein.  
XX  
XX Tumour suppressor activated pathway; tumour suppressor inhibited pathway;  
XX TSAP; TSIP; HUMSIAH; Human Homologue of the Drosophila seven in absentia;  
XX apoptosis; p53; cancer; Alzheimer's disease; probe; amplification; PCR;  
XX primer; hybridisation; diagnosis.  
XX

OS Homo sapiens.  
XX WO9722695-A2.  
XX  
XX 26-JUN-1997.  
XX  
XX 20-DEC-1996; 96WO-FR002061.  
XX  
XX 20-DEC-1995; 95FR-00015146.  
XX 18-APR-1996; 96FR-00004853.  
XX (DAUS-) FOND DAUSSET-CEPH JEAN.  
XX  
XX Talerman A, Amson R, Cohen D;  
XX WPI; 1997-341686/31.  
XX N-PSDB; AAT64820.  
XX  
XX New genes activated or inhibited during apoptosis - useful for treatment  
XX of, and for assessing risk of developing, cancer and Alzheimer's disease.  
XX  
XX Claim 1; Page 29-30; 51pp; French.  
XX  
XX This is the amino acid sequence of the human tumour suppressor activated  
XX pathway 3 (TSAP3) protein. The sequences AAT64813-T64823 represent genes  
XX of the tumour suppressor activated pathway (TSAP 1-8, also human TSAP 3,  
XX also designated HUMSIAH, i.e. Human Homologue of the Drosophila seven in  
XX absentia gene) or of the tumour suppressor inhibited pathway (TSIP 1 and  
XX 2). Expression of TSAP genes is induced during apoptosis while that of  
XX TSIP is induced by tumour suppressors and inhibited during apoptosis,  
XX especially where apoptosis is induced by p53. The sequences, vectors  
XX containing them or compounds that induce their expression are useful for  
XX treatment of cancer and Alzheimer's disease. Fragments of the sequences  
XX are useful as probes and amplification primers to determine  
XX predisposition to these diseases by detecting abnormalities in the genes  
XX  
XX Sequence 282 AA;  
SQ

Query Match 92.4%; Score 1489; DB 2; Length 282;  
Best Local Similarity 96.8%; Pred. No. 7.5e-138;  
Matches 273; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 17 MSRQTATALTGTSTKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQOSGHL 76  
DB 1 MSRQTATALTGTSTKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQOSGHL 60

QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136  
DB 61 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120

QY 137 CEFRPYSCPCPGASCCKWQSLDAMVPHLMHQHKSITTLQGEDIVFLATDINLPQAVDWM 196  
DB 121 CEFRPYSCPCPGASCCKWQSLDAMVPHLMHQHKSITTLQGEDIVFLATDINLPQAVDWM 180

QY 197 MQSCFGFHFMLVLEKQKYDGHQOFFAIVOLIIGTRKQAEAFYRLINGHRRRLTWEATP 256  
DB 181 MQSCFGFHFMLVLEKQKYDGHQOFFAIVOLIIGTRKQAEAFYRLINGHRRRLTWEATP 240

QY 257 RSIHEGIATAIMNSDCLVFTDTSIAQLFAENGNGINVTISM 298  
DB 241 RSIHEGIATAIMNSDCLVFTDTSIAQLFAENGNGINVTISM 282

RESULT 10  
ABB84447  
ID ABB84447 standard; peptide; 324 AA.  
XX  
XX ABB84447;  
XX  
XX 08-NOV-2002 (first entry)  
XX Human Siah2 protein.  
XX

KW Human; Siah2; seven in absentia homologue; ubiquitination pathway; cancer; infertility; inflammation; Siah; cytostatic; antiinflammatory; neuroprotective; immunosuppressive; apoptosis; NF kappa B signaling; neurological disorder; protein co-ordinate data.

XX Homo sapiens.

OS

PH Key Location/Qualifiers

FT Domain 80..114

FT /note= "RING domain"

XX WO200262838-A1.

XX

XX 15-AUG-2002.

XX

XX 06-FEB-2002; 2002WO-AU000118.

XX

XX 06-FEB-2001; 2001AU-00002908.

XX

XX (MACC-) MACCALLUM CANCER INST PETER.

XX (SVIN-) ST VINCENTS INST MEDICAL RES.

XX

XX Bowtell DDL, House CM, Parker MW, Polekhina G;

XX WPI; 2002-657521/70.

XX

XX New binding domain of a seven in absentia homolog protein, useful for identifying agonists or antagonists of the domain, for treating or preventing e.g. cancer, inflammation, infertility or other related conditions, or apoptosis.

XX

XX Disclosure; Fig 1b; 153pp; English.

XX

XX This invention describes a novel binding domain or a portion of the binding domain of a seven in absentia homologue (Siah) protein. Siah is an important protein in the ubiquitination pathway and the binding domain is capable of binding substrates, co-factors and interactors of the Siah protein. The products of the invention have cytostatic, antiinflammatory, neuroprotective and immunosuppressive activity. The binding domain is useful for identifying agonists and antagonists to the domain. The antagonist, agonist or ligand of the binding domain of Siah may be used for treating or preventing cancer e.g. breast cancer, inflammation, infertility or other related conditions, a disease relating to abnormal protein degradation, a pathological immune response, a disease relating to apoptosis, a disease relating to NF kappa B signaling, or a neurological disorder. This sequence represents the human Siah2 protein described in the disclosure of the invention. Note: This sequence contains 3-D protein co-ordinate data

XX

SQ Sequence 324 AA;

Query Match 74.4%; Score 1198; DB 5; Length 324;

Best Local Similarity 77.3%; Pred. No. 4e-109;

Matches 221; Conservative 27; Mismatches 30; Indels 8; Gaps 2;

QY 21 TATALPTGTSKCPSPORVPALTG-----TTASNNDLASLFECPVCFDYVLPPILOQCS 73

DB 37 TISAAGPGSSAVPAAAVISGPGGGGAGVPSPQHHELTSLFECPVCFDYVLPPILOQCA 96

QY 74 GHLVCSNCRPKLTCCPTCRGPL-GSIRNLAMEKVANSVLPCKYASSGCEITLPHTEKAD 132

DB 97 GHLVCNQCRQLSCCPTCRGALTPTSIRNLAMEKVASAVLPCKYATGCSLTHHTEKPE 156

QY 133 HEELCEFRYPSCPCGASCKWQGSIDAVNPHLMHQHSITTLQGEDIVFLATDINLPQAV 192

DB 157 HEDICEIRYPSCPCGASCKWQGSIRNLMAMEKVANSVLPCKYASSGCEITLPHTEKAD 216

QY 193 DWVMQSCGFHMLVLEKQEKYDGHQFFAIVQLIGTRKQAEFAIRLELNGHRRLLTW 252

DB 217 DWVMQSCGFHMLVLEKQEKYEGHQFFAIVLLIGTRKQAEFAIRLELNGHRRLLTW 276

QY 253 EATPRSIHGTAIWNNSDCLVFDTAIAHLFPADNGNLGINVTISTC 298

Db 277 EATPRSIHGDVAAAIMNSDCLVFDTAIAHLFPADNGNLGINVTISTC 322

RESULT 11

ABB84453

XX ID ABB84453 standard; peptide; 324 AA.

XX

XX AC ABB84453;

XX

XX DT 08-NOV-2002 (first entry)

XX

XX DE Human hsSiah2 protein.

XX

XX hsSiah2; seven in absentia homologue; ubiquitination pathway; cancer; infertility; inflammation; Siah; cytostatic; antiinflammatory; neuroprotective; immunosuppressive; apoptosis; NF kappa B signaling; neurological disorder; protein co-ordinate data.

XX

XX Homo sapiens.

XX WO200262838-A1.

XX

XX 15-AUG-2002.

XX

XX 06-FEB-2002; 2002WO-AU000118.

XX

XX 06-FEB-2001; 2001AU-00002908.

XX

XX (MACC-) MACCALLUM CANCER INST PETER.

XX (SVIN-) ST VINCENTS INST MEDICAL RES.

XX

XX Bowtell DDL, House CM, Parker MW, Polekhina G;

XX WPI; 2002-657521/70.

XX

XX New binding domain of a seven in absentia homolog protein, useful for identifying agonists or antagonists of the domain, for treating or preventing e.g. cancer, inflammation, infertility or other related conditions, or apoptosis.

XX

XX Disclosure; Fig 2; 153pp; English.

XX

XX This invention describes a novel binding domain or a portion of the binding domain of a seven in absentia homologue (Siah) protein. Siah is an important protein in the ubiquitination pathway and the binding domain is capable of binding substrates, co-factors and interactors of the Siah protein. The products of the invention have antiinflammatory, cytostatic, neuroprotective and immunosuppressive activity. The binding domain is useful for identifying agonists and antagonists to the domain. The antagonist, agonist or ligand of the binding domain of Siah may be used for treating or preventing cancer e.g. breast cancer, inflammation, infertility or other related conditions, a disease relating to abnormal protein degradation, a pathological immune response, a disease relating to apoptosis, a disease relating to NF kappa B signaling, or a neurological disorder. This sequence represents a Siah-related protein, hsSiah2 described in the disclosure of the invention. Note: This sequence contains 3-D protein co-ordinate data

XX

SQ Sequence 324 AA;

Query Match 74.4%; Score 1198; DB 5; Length 324;

Best Local Similarity 77.3%; Pred. No. 4e-109;

Matches 221; Conservative 27; Mismatches 30; Indels 8; Gaps 2;

QY 21 TATALPTGTSKCPSPORVPALTG-----TTASNNDLASLFECPVCFDYVLPPILOQCS 73

DB 37 TISAAGPGSSAVPAAAVISGPGGGGAGVPSPQHHELTSLFECPVCFDYVLPPILOQCA 96

QY 74 GHLVCSNCRPKLTCCPTCRGPL-GSIRNLAMEKVANSVLPCKYASSGCEITLPHTEKAD 132

DB 97 GHLVCNQCRQLSCCPTCRGALTPTSIRNLAMEKVASAVLPCKYATGCSLTHHTEKPE 156

QY 133 HEELCEFRYPSCPCGASCKWQGSIDAVNPHLMHQHSITTLQGEDIVFLATDINLPQAV 192

```

Db 157 HEDICEYRPSYPCPCGASCKWQSLDAVMPHLMHOKSITTLQGEDIVFLATDINLPGAV 216
QY 193 DVVMQSCFGFHMLVLEKQEKYDGHQOFFFAIVQLIGTRKQAFNPAFLNGLNRRRLTW 252
Db 217 DVVMQSCFGHHFMLVLEKQEKYEGHQOFFFAIVLLIGTRKQAFNPAFLNGLNRRRLTW 276
QY 253 EATPRSIHGIATATMNSDCLVFDTSIAQLFAENGLNINVTISM 298
Db 277 EATPRSIHGDVAAAIMNSDCLVFDTAIAHLFADNGLNINVTISTC 322

```

```

RESULT 12
ABR47593
ID ABR47593 standard; protein; 324 AA.
XX
AC ABR47593;
XX
DT 12-JUN-2003 (first entry)
XX
DE Breast cancer associated protein sequence SEQ ID NO:425.
XX
KW Human; breast cancer; cytostatic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2003004989-A2.
XX
PD 16-JAN-2003.
XX
PF 21-JUN-2002; 2002WO-US019669.
XX
PR 21-JUN-2001; 2001US-0299887P.
XX
PR 27-JUN-2001; 2001US-0301572P.
XX
PR 18-JUL-2001; 2001US-0306501P.
XX
PR 25-SEP-2001; 2001US-0325002P.
XX
PR 05-MAR-2002; 2002US-0362585P.
XX
PR 14-MAY-2002; 2002US-0380391P.
XX
PA (MILL-) MILLENIUM PHARM INC.
XX
PI Lillie J, Gannavarapu M, Glatt K, Hoersht S, Kamatkar S;
PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
PI Bast RC, Hortobagyi GN, Pusztai L, Mexic F, Sahin A, Mills GB;
XX
XX WPI; 2003-210381/20.
XX
XX N-PSDB; ACC50292.
XX
XX Breast cancer diagnosis or treatment by comparing the level of expression
XX of a marker in a patient sample with that in the control non-breast
XX cancer sample.
XX
XX Claim 1; SEQ ID NO 425; 128pp; English.
XX
XX
XX The present invention describes a method for assessing whether a patient
XX is afflicted with breast cancer. The method comprises comparing the level
XX of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
XX ABR47386 to ABR47632) in a patient sample and the normal level of
XX expression of the marker in a control non-breast cancer sample, where a
XX significant increase in the level of expression of the marker in the
XX patient sample and the normal level is an indication that the patient is
XX afflicted with breast cancer. The breast cancer associated sequences from
XX the present invention have cytostatic activities and can be used in gene
XX therapy. The method is useful for diagnosing and treating breast cancer.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 324 AA;
XX
XX Query Match 74.4%; Score 1198; DB 6; Length 324;
XX Best Local Similarity 77.3%; Pred. No. 4e-109;
XX Matches 221; Conservative 27; Mismatches 30; Indels 8; Gaps 2;

```

```

QY 21 TATALPTGTSCPPSQRPVATG-----TTASNNDLASLFECPVCFDYVLPPILOQCS 73
Db 37 TISAAGSSAVPAAAVISGPGGGGAGPVSPOHHELTSLFECPVCFDYVLPPILOQCA 96
QY 74 GHLVCSNCRPKLTCCPTCRGPL-GSIRNLAMEKVANSVLPCKYASSGCEITLPHTEKAD 132
Db 97 GHLVCNQCRQKLSCCPTCRGALTPTSIRNLAMEKVASAVLPCKYATTGCSLTLHTEKPE 156
QY 133 HEELCEPRPSPCPGASCKWQSLDAVMPHLMHOKSITTLQGEDIVFLATDINLPGAV 192
Db 157 HEDICEYRPSYPCPCGASCKWQSLDAVMPHLMHOKSITTLQGEDIVFLATDINLPGAV 216
QY 193 DVVMQSCFGFHMLVLEKQEKYDGHQOFFFAIVQLIGTRKQAFNPAFLNGLNRRRLTW 252
Db 217 DVVMQSCFGHHFMLVLEKQEKYEGHQOFFFAIVLLIGTRKQAFNPAFLNGLNRRRLTW 276
QY 253 EATPRSIHGIATATMNSDCLVFDTSIAQLFAENGLNINVTISM 298
Db 277 EATPRSIHGDVAAAIMNSDCLVFDTAIAHLFADNGLNINVTISTC 322

```

```

RESULT 13
ADRI4123
ID ADRI4123 standard; protein; 324 AA.
XX
AC ADRI4123;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human NF-kappaB pathway-associated protein SeqID124.
XX
KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
KW immunosuppressive; vulnary; gene therapy; immune disorder;
KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
KW viral replication; host cell survival; evasion of immune response;
KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW autoimmune disorder; hyper immune activity;
KW aberrant acute phase response; hypercongenital condition; birth defect;
KW necrotic lesion; wound; organ transplant rejection;
KW aberrant signal transduction; proliferating disorder; cancer;
KW HIV propagation; human.
XX
XX Homo sapiens.
XX
XX WO2004065577-A2.
XX
XX 05-AUG-2004.
XX
XX 13-JAN-2004; 2004WO-US000799.
XX
XX 14-JAN-2003; 2003US-0440068P.
XX
XX 12-MAY-2003; 2003US-0469757P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Nadler SG, Neubauer MG, Feder JN, Carman J;
XX
XX WPI; 2004-562168/54.
XX
XX N-PSDB; ADRI4122.
XX
XX New isolated polynucleotides and polypeptides associated with NF-kappaB
XX pathway, useful for diagnosing, treating, or preventing disorders or
XX diseases associated with NF-kappaB pathway.
XX
XX Claim 6; SEQ ID NO 124; 237pp; English.

```



Db 217 DWVMQSCFHHFVLEKQEKYEQHQFFAIVLLIGTRKQAFYRLNNGRRRLTW 276  
QY 253 EATPRSIHEGIATAMNSDCLVFDTSIAQLFAENGLGINVTISM 298  
Db 277 EATPRSIHDGVAAMNSDCLVFDTAIAHLFADNGLGINVTISTC 322

RESULT 15  
ID ABB84456 standard; peptide; 325 AA.  
XX AC ABB84456;  
DT 08-NOV-2002 (first entry)  
XX DE Murine mmsiah2 protein.  
XX KW mmsiah2; seven in absentia homologue; ubiquitination pathway; cancer;  
KW infertility; inflammation; Siah; cytostatic; antiinflammatory;  
KW neuroprotective; immunosuppressive; apoptosis; NF kappa B signaling;  
KW neurological disorder; protein co-ordinate data.

OS Mus musculus.  
XX WO200262838-A1.  
XX PN 15-AUG-2002.  
XX PF 06-FEB-2002; 2002WO-AU000118.  
XX PR 06-FEB-2001; 2001AU-00002908.  
XX PA (MACC-) MACCALLUM CANCER INST PETER.  
XX PA (SVIN-) ST VINCENTS INST MEDICAL RES.  
XX PI Bowtell DDL, House CM, Parker MW, Polekhina G;  
XX WPI; 2002-657521/70.

XX New binding domain of a seven in absentia homolog protein, useful for  
XX identifying agonists or antagonists of the domain, for treating or  
XX preventing e.g. cancer, inflammation, infertility or other related  
XX conditions, or apoptosis.  
XX Disclosure; Fig 2; 153pp; English.

XX This invention describes a novel binding domain or a portion of the  
XX binding domain of a seven in absentia homologue (Siah) protein. Siah is  
XX an important protein in the ubiquitination pathway and the binding domain  
XX is capable of binding substrates, co-factors and interactors of the Siah  
XX protein. The products of the invention have antiinflammatory, cytostatic,  
XX neuroprotective and immunosuppressive activity. The binding domain is  
XX useful for identifying agonists and antagonists to the domain. The  
XX antagonist, agonist or ligand of the binding domain of Siah may be used  
XX for treating or preventing cancer e.g. breast cancer, inflammation,  
XX infertility or other related conditions, a disease relating to abnormal  
XX protein degradation, a pathological immune response, a disease relating  
XX to apoptosis, a disease relating to NF kappa B signaling, or a  
XX neurological disorder. This sequence represents a Siah-related protein,  
XX mmsiah2 described in the disclosure of the invention. Note: This sequence  
XX contains 3-D protein co-ordinate data

SQ Sequence 325 AA;  
Query Match 73.9%; Score 1190; DB 5; Length 325;  
Best Local Similarity 76.9%; Pred. No. 2.5e-108;  
Matches 220; Conservative 27; Mismatches 31; Indels 8; Gaps 3;  
QY 21 TATLPTGTSKCPSPQSV---PALTG---TTASNNDLASLFECPCVCFDVLPPILQCO 73  
Db 38 TISAAGFGSSAPAAAVISGPGAGGADPVSPQHHELTSLFECPCVCFDVLPPILQCOA 97

QY 74 GHLVCSNCRPKLTCCPTCRGEL-GSIRNLAMEKVANSVLFPCKYASSGCEITLPHTEKAD 132  
Db 98 GHLVCSNCRPKLTCCPTCRGEL-GSIRNLAMEKVANSVLFPCKYATTGCSLTLLHTEKPE 157  
QY 133 HEELCEPRPYSCPCPGASCKWQGSIDAVMPLMHQHSITTLQGEDIVFLATDINLPGAV 192  
Db 158 HEDICEYRYPYSCPCPGASCKWQGSLEAVMSHLMHAHKSITTLQGEETVFLATDINLPGAV 217  
QY 193 DWVMQSCFHHFVLEKQEKYEQHQFFAIVLLIGTRKQAFYRLNNGRRRLTW 252  
Db 218 DWVMQSCFHHFVLEKQEKYEQHQFFAIVLLIGTRKQAFYRLNNGRRRLTW 277  
QY 253 EATPRSIHEGIATAMNSDCLVFDTSIAQLFAENGLGINVTISM 298  
Db 278 EATPRSIHDGVAAMNSDCLVFDTAIAHLFADNGLGINVTISTC 323

Search completed: April 25, 2005, 06:37:16  
Job time : 86 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 05:30:15 ; Search time 24 Seconds  
(without alignments)  
1194.691 Million cell updates/sec

Title: US-10-679-246-2

Perfect score: 1611

Sequence: 1 MWIIIFLLPPYVIFSEMSRQ.....IAQLFAENGNGINVTISMC 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79.\*

2: PIR1.\*

3: PIR2.\*

4: PIR3.\*

5: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1528	94.8	282	2	148763
2	1498	93.0	282	2	335754
3	1190	73.9	325	2	148765
4	1153.5	71.6	314	2	148195
5	1149	71.3	314	2	441544
6	1037.5	64.4	339	2	737470
7	439	27.3	315	2	750562
8	434.5	27.0	308	2	746026
9	429	26.6	327	2	709027
10	422	26.2	305	2	884848
11	419	26.0	315	2	747971
12	417	25.9	315	2	750561
13	385.5	23.9	336	2	750560
14	232.5	14.4	303	2	696692
15	215.5	13.4	313	2	896692
16	182.5	11.3	366	2	896692
17	120.5	7.5	458	2	829361
18	109	6.8	113	2	701657
19	107.5	6.7	468	2	748615
20	106	6.6	522	2	871821
21	105	6.5	89	2	703072
22	104	6.5	574	2	828275
23	104	6.5	974	2	888549
24	103.5	6.4	897	2	737813
25	101	6.3	587	2	724103
26	100	6.2	374	2	732286
27	99.5	6.2	298	2	707568
28	99	6.1	184	2	703178
29	98	6.1	292	2	888072

RESULT 1  
148763  
siab-1A protein - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I48763; S35753  
R;Della, N.G.; Senior, P.V.; Bowtell, D.D.  
Development 117, 1333-1343, 1993  
A;Title: Isolation and characterisation of murine homologues of the Drosophila seven in seven protein  
A;Reference number: I48763; MUID:94008536; PMID:8404535  
A;Accession: I48763  
A;Status: preliminary; translated from GE/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-282 <RES>  
A;Cross-references: UNIPROT:P61092; EMBL:Z19579; NID:9297034; PIDN:CAA79630.1; PID:G297030;  
C;Superfamily: Drosophila developmental protein sina; RING finger homology

Query Match 94.8%; Score 1528; DB 2; Length 282;  
Best Local Similarity 99.6%; Pred. No. 2.9e-123;  
Matches 281; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 17 MSRQTATLPTGSKCPSPSRVPAITGTTASNNDLASLFEPCVCFDYLPPILQCSGHL 76  
Db 1 MSRQTATLPTGSKCPSPSRVPAITGTTASNNDLASLFEPCVCFDYLPPILQCSGHL 60  
Qy 77 VCSNCRPKLTCCPTCRGPLGSIENLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136  
Db 61 VCSNCRPKLTCCPTCRGPLGSIENLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120  
Qy 137 CEFPRYSCPCPGASCKWQSGSLDAVMPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 196  
Db 121 CEFPRYSCPCPGASCKWQSGSLDAVMPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 180  
Qy 197 MQSCFGFHFMLVEKEKYGCHQOFFFAIVOLIGTRKQAEFPAYRLELNHRRRUTWTATP 256  
Db 181 MQSCFGFHFMLVEKEKYGCHQOFFFAIVOLIGTRKQAEFPAYRLELNHRRRUTWTATP 240  
Qy 257 RSTHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISMC 298  
Db 241 RSTHEGIATAMNSDCLVFDTSIAQLFAENGNGINVTISMC 282

RESULT 2  
335754  
siab-1B protein - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: I48764; S35754  
R;Della, N.G.; Senior, P.V.; Bowtell, D.D.  
Development 117, 1333-1343, 1993  
A;Title: Isolation and characterisation of murine homologues of the Drosophila seven in seven protein  
A;Reference number: I48763; MUID:94008536; PMID:8404535

ALIGNMENTS

30 97.5 6.1 826 2 A60385 monocyte surface a  
31 97 6.0 235 2 T19328 hypothetical prote  
32 97 6.0 421 2 S44864 R05D3.3 protein -  
33 96.5 6.0 326 2 S34427 tristetrapoline p  
34 96 6.0 425 2 A84849 probable RING zinc  
35 94.5 5.9 562 2 H88071 protein ZK1240.3 [  
36 94 5.8 2447 2 T16870 hypothetical prote  
37 93 5.8 421 2 A56550 Krox-20 - African  
38 93 5.8 2652 1 VFIH2 genome polyprotein  
39 93 5.8 3051 2 S42373 hypothetical prote  
40 91.5 5.7 255 2 H96776 hypothetical prote  
41 91.5 5.7 1589 2 T42233 submaxillary mucin  
42 91 5.6 226 2 E86326 protein F18014.3 [  
43 91 5.6 442 2 A57041 transcription regu  
44 90.5 5.6 563 2 A36054 mucin homolog - bo  
45 90.5 5.6 591 2 I48141 acrogranin - guine



```
A;Accession: I48764
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-282 <RES>
C;Cross-references: UNIPROT:Q06985; EMBL:Z19580; NID:9297801; PIDN:CAA79631.1; PID:92978
C;Superfamily: Drosophila developmental protein sina; RING finger homology

Query Match 93.0%; Score 1498; DB 2; Length 282;
Best Local Similarity 97.5%; Pred. No. 1.1e-120; Indels 0; Gaps 0;
Matches 275; Conservative 4; Mismatches 3;

Qy 17 MSRTATALPTGTSKCPSPQSRVPAITTTASNNDLASLFECPVCFDYVLPPIILQCSGHL 76
Db 1 MSRQAATALGTGTSKCPSPQSRVPAITTTASNNDLASLFECPVCFDYVLPPIILQCSGHL 60

Qy 77 VCSNCRPKLTCCPTCRGPGLSIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136
Db 61 VCSNCRPKLTCCPTCRGPGLSIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120

Qy 137 CERFPYSCPCPGASCKWQSGISDAVPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 196
Db 121 CERFPYSCPCPGASCKWQSGISDAVPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 180

Qy 197 MQSCFGFHFHMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFAYRLELNGHRRRLTWEATP 256
Db 181 MQSCFGFHFHMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFAYRLELNGHRRRLTWEATP 240

Qy 257 RSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 298
Db 241 RSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 282

RESULT 3
I48765
siah-2 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48765; S35755
R;Della, N.G.; Senior, P.V.; Bowtell, D.D.
Development 117, 1333-1343, 1993
A;Title: Isolation and characterisation of murine homologues of the Drosophila seven in
A;Reference number: I48763; MUID:94008536; PMID:8404535
A;Accession: I48765
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-325 <RES>
A;Cross-references: UNIPROT:Q06986; EMBL:Z19581; NID:9297036; PIDN:CAA79632.1; PID:92970
C;Superfamily: Drosophila developmental protein sina; RING finger homology

Query Match 73.9%; Score 1190; DB 2; Length 325;
Best Local Similarity 76.9%; Pred. No. 2.8e-94;
Matches 220; Conservative 27; Mismatches 31; Indels 8; Gaps 3;

Qy 21 TATALPTGTSKCPSPQSRV---PALTG----TTASNNDLASLFECPVCFDYVLPPIILQCS 73
Db 38 TISAAGPSSAVPAAAVISPGAGGADPVSPQHSLTSLFECPVCFDYVLPPIILQCS 97

Qy 74 GHLVCSNCRPKLTCCPTCRGPL-GSIRNLAMEKVANSVLPCKYASSGCEITLPHTEKAD 132
Db 98 GHLVCSNCRPKLTCCPTCRGALTFSIRNLAMEKVANSVLPCKYASSGCEITLPHTEKPE 157

Qy 133 HEELCEFRPYSCPCPGASCKWQSGISDAVPHLMHQHSITTLQGEDIVFLATDINLPGAV 192
Db 158 HEDICEFRPYSCPCPGASCKWQSGISDAVPHLMHQHSITTLQGEIVFLATDINLPGAV 217

Qy 193 DWMVMSQSCFGFHFHMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFAYRLELNGHRRRLTW 252
Db 218 DWMVMSQSCFGFHFHMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFAYRLELNGHRRRLTW 277

Qy 253 EATPRSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 298
Db 278 EATPRSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 323
```

```
RESULT 4
A36195
developmental protein sina - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 14-Dec-1990 #sequence_revision 13-Jan-1993 #text_change 09-Jul-2004
C;Accession: A36195
R;Carthew, R.W.; Rubin, G.M.
Cell 63, 561-577, 1990
A;Title: seven in absentia, a gene required for specification of R7 cell fate in the Dro
A;Reference number: A36195; MUID:91029488; PMID:2146028
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-314 <CAR>
A;Cross-references: UNIPROT:P21461; GB:M38384; NID:9158466; PID:9158467
C;Genetics:
A;Gene: FlyBase:sina
A;Cross-references: FlyBase:FBgn0003410
C;Superfamily: Drosophila developmental protein sina; RING finger homology
C;Keywords: DNA binding; nucleus; zinc finger

Query Match 71.6%; Score 1153.5; DB 2; Length 314;
Best Local Similarity 75.3%; Pred. No. 3.6e-91;
Matches 217; Conservative 18; Mismatches 42; Indels 11; Gaps 1;

Qy 21 TATALPTG-----TSKCPSPQSRVPAITTTASNNDLASLFECPVCFDYVLPPIIL 69
Db 26 TWTSTGTGSSAGTSSANTSSSSSSLSAGGAGMSADLTSLFECPVCFDYVLPPIIL 85

Qy 70 QCQSGHLVCSNCRPKLTCCPTCRGPGLSIRNLAMEKVANSVLPCKYASSGCEITLPHTE 129
Db 86 QCQSGHLVCSNCRPKLTCCPTCRGPGLSIRNLAMEKVANSVLPCKYASSGCEITLPHTE 145

Qy 130 KADHEELCEFRPYSCPCPGASCKWQSGISDAVPHLMHQHSITTLQGEDIVFLATDINLP 189
Db 146 KTEHEETCECRPYLPCPGASCKWQSGISDAVPHLMHQHSITTLQGEDIVFLATDINLP 205

Qy 190 GAVDWMVMSQSCFGFHFHMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFAYRLELNGHRR 249
Db 206 GAVDWMVMSQSCFGFHFHMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFAYRLELNGHRR 265

Qy 250 LTWEATPRSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 297
Db 266 LTWEATPRSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 313

RESULT 5
A41544
developmental protein sina - fruit fly (Drosophila virilis)
C;Species: Drosophila virilis
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: A41544
R;Neufeld, T.P.; Carthew, R.W.; Rubin, G.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10203-10207, 1991
A;Title: Evolution of gene position: chromosomal arrangement and sequence comparison of t
A;Reference number: A41544; MUID:92052239; PMID:1946441
A;Accession: A41544
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-314 <NEU>
A;Cross-references: UNIPROT:P29304; GB:M77281
C;Genetics:
A;Gene: FlyBase:Dvir/sina
A;Cross-references: FlyBase:FBgn0013142
C;Superfamily: Drosophila developmental protein sina; RING finger homology
C;Keywords: zinc finger

Query Match 71.3%; Score 1149; DB 2; Length 314;
Best Local Similarity 75.3%; Pred. No. 8.7e-91;
Matches 216; Conservative 19; Mismatches 42; Indels 10; Gaps 1;

Qy 21 TATALPTG-----TSKCPSPQSRVPAITTTASNNDLASLFECPVCFDYVLPPIILQ 70
```

Db 27 TINTSSGTSSAGNTSSANTSSSSSSSSSSAGGGAGSADTSLFECVCFDYVLPPILO 86  
Qy 71 COSGHLVCSNCRPKLTCCPTCRGPLGSIIRNLAMEKVANSVLPCKYASSGCEITLPHTEK 130  
Db 87 CSSGHLVCSNCRPKLTCCPTCRGPLANIRNLAMEEASNVKFPCKHSGYGCTASLAVYTEK 146  
Qy 131 ADHEELCEFRPYSCPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLP 190  
Db 147 TEHEETCECRPLPCPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLP 206  
Qy 191 AVDWMMQSCFGFHMVLVLEKQEKYDGHQOQFAIVQLIGTRKQAEINPAYRLELNGHRRRL 250  
Db 207 AVDWMMQSCFGFHMVLVLEKQEKYDGHQOQFAIVQLIGTRKQAEINPAYRLELNGHRRRL 266  
Qy 251 TWEATPRSIHEGIATAMNSDCLVFTDSIAQLFAENGNLGINVTISM 297  
Db 267 TWEAMPRSIHEGVASAIHNSDCLVFTDSIAQLFADNGLNINVTISL 313

## RESULT 6

T37470  
seven-in-absentia protein homolog-1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: T37470  
R:Holloway, A.J.; Bowtell, D.D.L.  
submitted to the EMBL Data Library, February 1997  
A:Description: Cloning and characterisation of the Caenorhabditis elegans homologue of  
A:Reference number: 221707  
A:Accession: T37470  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-339 <HOL>  
A:Cross-references: UNIPROT:Q965X6; EMBL:U89792; PIDN:AAB94380.1  
C:Genetics:  
A:Gene: sia-1  
A:Introns: 129/1; 321/1  
C:Superfamily: Drosophila developmental protein sina; RING finger homology

Query Match 64.4%; Score 1037.5; DB 2; Length 339;  
Best Local Similarity 71.5%; Pred. No. 3.3e-81;  
Matches 183; Conservative 31; Mismatches 39; Indels 3; Gaps 1;  
Qy 44 TTASNNDLASFECVCFDYVLPPILOQSGHLVCSNCRPKLTCCPTCRGPLGSIIRNLAM 103  
Db 77 TDDSSAELSIFECVCFDYVLPPILOQSGHLVCSNCRPKLTCCPTCRGPTSPVNLGL 136  
Qy 104 EKVANSVLPCKYASSGCEITLPHTEKADHEELCEFRPYSCPCGASCKWQSLDVAVPH 163  
Db 137 EKIANTVRFPCFKSTGSCPLNFHFKADKTEHEELCEFRPYSCPCGASCKWQSLDVAVPH 196  
Qy 164 LMHQHKSITTLQGEDIVFLATDINLPAGVMMQSCFGFHMVLVLEKQEKYD--GHQ 220  
Db 197 LKXIKHSITTLQGEDIVFLATDINLPAGVMMQSCFGFHMVLVLEKQEKYDPAQTOM 256  
Qy 221 FFAIVQLIGTRKQAEINPAYRLELNGHRRRLTWEATPRSIHEGIATAMNSDCLVFTDSIA 280  
Db 257 FYAVVQLIGSKKADNFVYRLSLSRRMSWEATPRSIHEGVVVAQQSDCLAFDSNAA 316  
Qy 281 QLFPAENGNLGINVTIS 296  
Db 317 QLFPAENGNLGINVTIS 332

## RESULT 7

T50562  
SINA2 protein [imported] - Vitis vinifera  
C:Species: Vitis vinifera  
C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 09-Jul-2004  
C:Accession: T50562  
R:Brehm, I.; Korfei, M.; Preisig-Mueller, R.; Kindl, H.  
submitted to the EMBL Data Library, November 1998

A:Description: A nuclear localized zinc finger protein found in a plant is homologous to  
A:Reference number: Z25132  
A:Accession: T50562  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-315 <BRE>  
A:Cross-references: UNIPROT:Q9XGC3; EMBL:Y18472; PIDN:CAB40578.1  
C:Superfamily: Drosophila developmental protein sina; RING finger homology

Query Match 27.3%; Score 439; DB 2; Length 315;  
Best Local Similarity 32.7%; Pred. No. 4.8e-30;  
Matches 98; Conservative 45; Mismatches 121; Indels 36; Gaps 7;

Qy 18 SROTAT----ALPTGTSCKPPSQRPVLPALGT--TASNNDLASFECVCFDYVLPPILO 71  
Db 14 SHSTVTVDYDIAMTKADTNLTAKTGLSGKGGISSPVGVELLECPVCTSLMYPPIYQC 73  
Qy 72 QSGHLVCSNCRPKL-TCCPTCRGPLGSIIRNLAMEKVANSVLPCKYASSGCEITLPHTEK 130  
Db 74 PSHTLCSNCKSRVHNCCTCRHELGDICLALEKVAESLELPCRYQSLGCHDIFPYYSK 133  
Qy 131 ADHEELCEFRPYSCPCGASCKWQSLDVAVPHLMHQHKSITTLQ 176  
Db 134 LKHEQOQCRFPYNCYPAGFECSTGDIPTLVEHLKGDHKVDMHDGCTFNHRYVKSNPQEV 193  
Qy 177 EDIVFLATDINLPAGVMMQSCFGFHMVLVLEKQEKYDGHQOQFAIVQLIGTRKQAEIN 236  
Db 194 ENATWMLTVFN-----CFGKQFCLHPEAFQ-LGTAPVYMAFLRPMGMDNEAKK 240  
Qy 237 FAYRLELNGHRRRLTWEATPRSIHEGIATAMNSDCLVFTDSIAQLFA--ENGNLGINVT 294  
Db 241 FSYSLVGGNSRKLIMQGVPRISDRSHKVRDQSGLIQRIQNLALYFSGDQRLKLRVT 300

## RESULT 8

T46026  
hypothetical protein T10K17.250 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C:Accession: T46026  
R:Benes, V.; Wurmbach, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z23019  
A:Accession: T46026  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-308 <BEN>  
A:Cross-references: UNIPROT:Q9M2P4; EMBL:AL132977  
A:Experimental source: cultivar Columbia; BAC clone T10K17  
C:Genetics:  
A:Map position: 3  
A:Introns: 73/3; 202/3  
A:Note: T10K17.250  
C:Superfamily: Drosophila developmental protein sina; RING finger homology

Query Match 27.0%; Score 434.5; DB 2; Length 308;  
Best Local Similarity 31.7%; Pred. No. 1.1e-29;  
Matches 97; Conservative 47; Mismatches 115; Indels 47; Gaps 7;

Qy 18 SROTATLPTGTSCKPPSQRPVLPALGTGTTA-----SNNDLASFECVCFDYVLPPILO 70  
Db 14 SNSTGMDYEVKTAKVEYNNNNKPTKPSAGIGKYGIIHNSNGVYELLECPVCTNLMLYPPIHQ 73  
Qy 71 COSGHLVCSNCRPKL-TCCPTCRGPLGSIIRNLAMEKVANSVLPCKYASSGCEITLPHTE 129  
Db 74 CPNGHLCSNCKURVQNTCTCRYELGNIRKLALEKVAESLEVPCTQNLGCHDIFPYYS 133  
Qy 130 KADHEELCEFRPYSCPCGASCKWQSLDVAVPHLMHQHKSITTLQ 175  
Db 134 KLKHEQHCRFPYTCYPVAGSECSVTGDIPTLVHLKDDHKVDMHDGCTFNHRYVKSNPHE 193  
Qy 176 GEDIVFLATDINLPAGVMMQSCFGFHMVLVLEKQEKYDGHQ-----OFFAIVQLIGT 230

Db 194 VENATWMLTVFN-----CFGRQCL-----HFEAFQLGMAPVYMAFLRFMGD 235

Qy 231 RKOENPAYELENGHRRRLTWTATPSRSHGATATMNSDCLVFTSIAQLFA--ENG 288

Db 236 ENEAKKPSYSLEVGAGHKLTKGIPRSIRDSHRKVRDSQDGLIIPRNALALYFSGGDR 295

Qy 289 LGINVT 294

Db 296 LKLRVT 301

RESULT 9

hypothetical protein T27E11.120 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C:Accession: T09027

R:Bayan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16533

A:Accession: T09027

A:Molecule type: DNA

A:Residues: 1-327 <BE>

A:Cross-references: UNIPROT:Q9STN8; EMBL:AL078579; GSPDB:GN00062; ATSP:T27E11.120

A:Experimental source: cultivar Columbia; BAC clone T27E11

C:Genetics:

A:Gene: ATSP:T27E11.120

A:Map position: 4

A:Introns: 77/3; 206/3

C:Superfamily: Drosophila developmental protein sina; RING finger homology

Query Match 26.6%; Score 429; DB 2; Length 327;

Best Local Similarity 35.1%; Pred. No. 3.6e-29;

Matches 102; Conservative 45; Mismatches 126; Indels 18; Gaps 8;

Qy 11 YVFISEMSROTATLPTGTSKCPSPORVPALTGTATSNNDLASLFEPCVCFDYVLPILQ 70

Db 26 YQFSSTKTHGGAATAVTVNGPTATAPA-----TSVYELLECPVCTYSMPPIHQ 77

Qy 71 COSGHLVCSNCRPKL--TCCTPCRGPLGSIRNLAMEKVANSVLPFCYKASSGCEITLPHTE 129

Db 78 CHNGHTLCTCKVRVHNRCPCTCRQELGDIRCLALEKVAESLELPCFVNLGCPPIFYIS 137

Qy 130 KADHEELCEFRYPSCPCGASCKWQGLSDAVMPLMHQHSITTLQGEDIVFLATDINLP 189

Db 138 KKHESLGNFRYPSCYAGSECGIVGDIPLVAHLRDDHK--VDMHAGSTENHRYVKS-N-P 195

Qy 190 GAVD---WVM-MQSCFGFHFMLVLEKQBYDHOQFFAIVOLIGTRKQAFNFAIRLELNG 245

Db 196 REVENATWMLTVFHCQGVYFCLHPEAFQLGGMG-PVYMAFLRFMGDEEDARSYSLEYVG 254

Qy 246 HRRRLTWTATPSRSHGATATMNSDCLVFTSIAQLFA--ENG 294

Db 255 SGKLTWETGTPRSIRDSHRKVRDSNDGLIIOQNMAFLFFSGGDRKELKURVT 305

RESULT 10

probable RING zinc finger protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: E84848

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umavam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eissen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84848

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-305 <STO>

A:Cross-references: UNIPROT:P93748; GB:AE002093; NID:g1871185; PIDN:AAB63545.1; GSPDB:GN

C:Genetics:

A:Gene: At2g41980

A:Map position: 2

C:Superfamily: Drosophila developmental protein sina; RING finger homology

Query Match 26.2%; Score 422; DB 2; Length 305;

Best Local Similarity 31.2%; Pred. No. 1.3e-28;

Matches 94; Conservative 51; Mismatches 108; Indels 48; Gaps 7;

Qy 16 EMSROTATLPTGTSKCPSPORVPALTGTATSNNDLASLFEPCVCFDYVLPILQCSGH 75

Db 24 KMAKVEANSKPTKSGS-----SGIKFHSNGVYELLECPCVTNLMYPIHQCPNGH 75

Qy 76 LVCSNCRPKL--TCCTPCRGPLGSIRNLAMEKVANSVLPFCYKASSGCEITLPHTEKADHE 134

Db 76 TLCSCKLVRQNTCTPCRYELGNIRCLALEKVAESLEVPYQNLGCDQIFPYSKLKE 135

Qy 135 ELCEFRYPSCPCGASCKWQGLSDAVMPLMHQHK-----SITTLQGEDIV 180

Db 136 QHCRFRASCPYAGSECSVTGDIPTLVDLHKDDHMDHGGCTFNHRYVKSHPHEVAT 195

Qy 181 ELATDINLPGAVDWVMQSCFGFHFMLVLEKQEKYDGHQ-----OFFAIVQLIGTRKQAE 235

Db 196 WMLTVFN-----CFGRQCL-----HFEAFQLGMAPVYMAFLRFMGDENEAK 237

Qy 236 NFAYRLLENGHRRRLTWTATPSRSHGATATMNSDCLVFTSIAQLP--AENG 293

Db 238 KFSYSLEVGAGHKLTKGIPRSIRDSHRKVRDSQDGLIIPRNALALYFSGSKELKRV 297

Qy 294 T 294

Db 298 T 298

RESULT 11

T47971

seven in absentia-like protein - Arabidopsis thaliana

N:Alternate names: protein F15G16.180

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004

C:Accession: T47971

R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetj

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z24480

A:Accession: T47971

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <DEH>

A:Cross-references: UNIPROT:O84JL3; EMBL:AL132959

A:Experimental source: cultivar Columbia; BAC clone F15G16

C:Genetics:

A:Map position: 3

A:Introns: 65/3; 194/3

A:Note: F15G16.180

C:Superfamily: Drosophila developmental protein sina; RING finger homology

Query Match 26.0%; Score 419; DB 2; Length 315;

Best Local Similarity 34.0%; Pred. No. 2.5e-26;

Matches 91; Conservative 46; Mismatches 109; Indels 22; Gaps 7;

Qy 40 ALTGTATSNNDLASLFEPCVCFDYVLPILQCSGHVLCNSCRPKL--TCCTPCRGPLGI 98

Db 35 AASGLPTTTSVHLELCEPCVTNSMYPIHQCHNGHTLCTCKARVHNRCPCTCRQELGDI 94

Qy 99 RNLAAMEKVANSVLPFCYKASSGCEITLPHTEKADHELCFEPYSCPCGASCKWQGLSD 158

Db 95 RCLALEKVAESLELPCCKHMSLGCPEIFPYSKLKHETVCNFRPYSCPYAGSECSVTGDI 154

Qy 159 AVMPHLMHQHK-----SITTLQGEDIVFLATDINLPGAVDWVM-MQSCFGFHFMLVLEKQ 213

Db 155 FLVAHLRDDHKVDMHSGCTFNHRYVKSHPREVE---NATWMLTVFHCQGVYFCL----- 205

Qy 214 KYDGHQ-----OFFAIVQLIGTRKQAFNFAIRLENGHRRRLTWTATPSRSHGATATM 268



C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: B96692  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-313 <STO>  
A/Cross-references: UNIPROT:Q9C6H3; GB:AEO05173; NID:G11054583; PIDN:AAG27858.1; GSPDB:G  
C/Genetics:  
A/Gene: T1217.7  
A/Map position: 1

Query Match 13.4%; Score 215.5; DB 2; Length 313;  
Best Local Similarity 30.6%; Pred. No. 6.3e-11;  
Matches 44; Conservative 23; Mismatches 64; Indels 13; Gaps 3;  
QY 36 QRVPAITGTTASNNDLA-----SLFECPCVCFDYVLPPILOQOSGHLVGSNCRPKL 85  
Db |||||:  
13 QRVFSSVESVGGDAVASGTLFELDLDDCPICCHALTSPIFQCDNGHIACSSCCTKL 72  
QY 86 -TCCPTCRGPIGSIIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHBEELCEFRFYSC 144  
Db |||||:  
73 RNKCPSCALPIGNFRSIRMERVVEAVVTCFNVKHGCTEKFSYKELIHEKDCRFA--LC 130  
QY 145 PCPGASCKWQGSILDVMPHLMHQH 168  
Db |||||:  
131 YCPAPNCNYSGVYKDYSHFYVNH 154

Search completed: April 25, 2005, 06:39:13  
Job time : 27 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 03:59:42 ; Search time 82 Seconds  
(without alignments)  
1860.971 Million cell updates/sec

Title: US-10-679-246-2

Perfect score: 1611

Sequence: 1 MWIIIFLLPPYFISEMSRQ.....IAQLFAENGNGINVTISM 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1532	95.1	282	1	SIH1_HUMAN
2	1529	94.9	282	2	Q6GQJ5
3	1528	94.8	282	1	SIH1_MOUSE
4	1528	94.8	282	1	SIH1_RAT
5	1508	93.6	282	1	SIH1_BRARE
6	1494	92.7	282	1	SIH1_MOUSE
7	1210	75.1	311	2	Q7Q520
8	1199	74.4	325	1	SIH2_MOUSE
9	1199	74.4	325	1	SIH2_RAT
10	1198	74.4	324	1	SIH2_HUMAN
11	1192.5	74.0	331	1	SIH2_BRARE
12	1177.5	73.1	313	1	SIH2_XENLA
13	1156.5	71.8	331	1	SINA_DROWI
14	1153.5	71.6	314	1	SINA_DROWI
15	1153.5	71.6	314	1	SINA_DROME
16	1149	71.3	314	1	SINA_DROWI
17	1068	66.3	200	2	Q8JH27
18	1042.5	64.7	371	1	SINA_SCHMA
19	1037.5	64.4	339	1	SIH1_CABEL
20	737	45.7	326	2	Q8I146
21	729.5	45.3	354	2	Q8I1H9
22	702.5	43.6	351	1	SINL_DROME
23	646.5	40.1	139	2	Q6X927
24	573	35.6	128	2	Q69DP6
25	571	35.4	128	2	Q69DN8
26	568	35.3	128	2	Q69DM8
27	568	35.3	128	2	Q69DN1
28	568	35.3	128	2	Q69DN2
29	568	35.3	128	2	Q69DN3
30	568	35.3	128	2	Q69DN7
31	568	35.3	128	2	Q69DP3

## RESULT 1

ID	SIH1_HUMAN	STANDARD;	PRT;	282 AA.
AC	Q8IUQ4; Q43269; Q92880;			
DT	05-JUL-2004 (Rel. 44, Created)			
DT	05-JUL-2004 (Rel. 44, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Ubiquitin ligase SIAH1 (EC 6.3.2.-) (Seven in absentia homolog 1) (Siah-1) (Siah-1a).			
GN	Name=SIAH1; Synonyms=HUMSIAH;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Intestinal epithelium;			
RX	MEDLINE=96392362; PubMed=8799150; DOI=10.1073/pnas.93.17.9039;			
RA	Nemani M., Linares-Cruz G., Bruzoni-Giovanelli H., Roperch J.-P.,			
RA	Tuynder M., Bougueleret L., Cherif D., Medhioub M., Pasturaud P.,			
RA	Alvaro V., Der Sarkissian H., Cazes L., Le Paslier D., Le Gall I.,			
RA	Israeli D., Dausset J., Sigaux F., Chumakov I., Oren M., Calvo F.,			
RA	Amson R.B., Cohen D., Telerman A.;			
RT	"Activation of the human homologue of the Drosophila sina gene in			
RT	apoptosis and tumor suppression."			
RL	Proc. Natl. Acad. Sci. U.S.A. 93:9039-9042(1996).			
[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 1), SUBCELLULAR LOCATION, AND TISSUE			
RC	SPECIFICITY.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=9806768; PubMed=9403064; DOI=10.1006/geno.1997.4997;			
RA	Hu G., Chung Y.-L., Glover T., Valentine V., Look A.T., Fearon E.R.;			
RT	"Characterization of human homologs of the Drosophila seven in			
RT	absentia (sina) gene."			
RL	Genomics 46:103-111(1997).			
[3]				
RP	SEQUENCE FROM N.A.			
RC	PubMed=10956387;			
RX	Medhioub M., Vaurty C., Hamelin R., Thomas G.;			
RA	"Lack of somatic mutation in the coding sequence of SIAH1 in tumors			
RT	hemizygous for this candidate tumor suppressor gene."			
RL	Int. J. Cancer 87:794-797(2000).			
[4]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Retina;			
RA	Koehler K., Beyer A., Meves H.-W., Weil B., Amid C., Osanger A.,			
RA	Fobo G., Han M., Wiemann S.;			
RL	Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.			
[5]				
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
RC	TISSUE=Brain, and Pancreas;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			

## ALIGNMENTS

32	568	35.3	128	2	Q69DP4
33	568	35.3	128	2	Q69DP5
34	568	35.3	128	2	Q69DP7
35	567	35.2	128	2	Q69DN6
36	567	35.2	128	2	Q69DN9
37	567	35.2	128	2	Q69DP2
38	552	34.3	125	2	Q69DN4
39	552	34.3	125	2	Q69DP1
40	547	34.0	124	2	Q69DN5
41	542	33.6	123	2	Q69DP0
42	533	33.1	123	2	Q69DN0
43	483	30.0	296	2	Q8IWO3
44	462.5	28.7	308	2	Q6Z6Z3
45	439	27.3	315	2	Q9XGC3

Q69dp4	drosophila
Q69dp5	drosophila
Q69dp7	drosophila
Q69dn6	drosophila
Q69dn9	drosophila
Q69dp2	drosophila
Q69dn4	drosophila
Q69dp1	drosophila
Q69dn5	drosophila
Q69dp0	drosophila
Q69dn0	drosophila
Q8IWO3	homo sapien
Q6Z6Z3	oryza sativ
Q9XGC3	vitis vinif

- RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RN FUNCTION IN DEGRADATION OF DCC, SUBCELLULAR LOCATION, AND INTERACTION  
RN WITH UBE21.  
RP PubMed=9334332;  
RX Hu G., Zhang S., Vidal M., Baer J.L., Xu T., Fearon E.R.;  
RT "Mammalian homologs of seven in absentia regulate DCC via the  
RT ubiquitin-proteasome pathway.";  
RL Genes Dev. 11:2701-2714(1997).  
RN [7]  
RN INTERACTION WITH BAG1, AND SUBCELLULAR LOCATION.  
RP PubMed=958267; DOI=10.1093/emboj/17.10.2736;  
RX Matsuzawa S., Takayama S., Froesch B.A., Zapata J.M., Reed J.C.;  
RA "p53-inducible human homologue of Drosophila seven in absentia (Siah)  
RT inhibits cell growth: suppression by BAG-1.";  
RL EMBO J. 17:2736-2747(1998).  
RN [8]  
RN FUNCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF GLU-40; CYS-41;  
RP CYS-44; CYS-55; HIS-59; ARG-66; LYS-68; ARG-76; HIS-152; HIS-202 AND  
RP LEU-211.  
RX PubMed=9858595;  
RA Hu G., Fearon E.R.;  
RT "Siah-1 N-terminal RING domain is required for proteolysis function,  
RT and C-terminal sequences regulate oligomerization and binding to  
RT target proteins.";  
RL Mol. Cell. Biol. 19:724-732(1999).  
RN [9]  
RN FUNCTION IN DEGRADATION OF KIF22, AND INTERACTION WITH ALPHA-TUBULIN.  
RP PubMed=11146551; DOI=10.1038/sj.onc.1204002;  
RX Germani A., Bruzzoni-Giovanelli H., Fellous A., Gisselbrecht S.,  
RA Varin-Blank N., Calvo F.;  
RT "Siah-1 interacts with alpha-tubulin and degrades the kinesin Kid by  
RT the proteasome pathway during mitosis.";  
RL Oncogene 19:5997-6006(2000).  
RN [10]  
RN FUNCTION IN DEGRADATION OF MYB.  
RP PubMed=10747903; DOI=10.1074/jbc.M000372200;  
RX Tanikawa J., Ichikawa-Iwata E., Kanai-Ishii C., Nakai A.,  
RA Matsuzawa S.-I., Reed J.C., Ishii S.;  
RT "p53 suppresses the c-Myc-induced activation of heat shock  
RT transcription factor 3.";  
RL J. Biol. Chem. 275:15578-15585(2000).  
RN [11]  
RN FUNCTION IN DEGRADATION OF CTNNB1, AND SUBUNIT OF A COMPLEX WITH  
RP UBE2D1; CACBP; SKP1A; APC AND TBL1X.  
RX PubMed=11389839; DOI=10.1016/S1097-2765(01)00242-8;  
RA Matsuzawa S.-I., Reed J.C.;  
RT "Siah-1, SIP, and Ebi collaborate in a novel pathway for beta-catenin  
RT degradation linked to p53 responses.";  
RL Mol. Cell. Biol. 21:915-926(2001).  
RN [12]  
RN FUNCTION IN DEGRADATION OF CTNNB1.  
RP PubMed=11389840; DOI=10.1016/S1097-2765(01)00241-6;  
RX Liu J., Stevens J., Rote C.A., Yost H.J., Hu Y., Neufeld K.L.,  
RA White R.L., Matsunami N.;  
RT "Siah-1 mediates a novel beta-catenin degradation pathway linking p53  
RT to the adenomatous polyposis coli protein.";  
RL Mol. Cell. Biol. 21:927-936(2001).  
RN [13]  
RN FUNCTION IN DEGRADATION OF POU2AF1, AND SUBCELLULAR LOCATION.  
RP PubMed=11483517; DOI=10.1093/emboj/20.15.4143;  
RX Tiedt R., Bartholdy B.A., Matthias G., Newell J.W., Matthias P.;  
RA "The RING finger protein Siah-1 regulates the level of the  
RT transcriptional coactivator OBF-1.";  
RL EMBO J. 20:4143-4152(2001).  
RN [14]  
RN FUNCTION IN DEGRADATION OF POU2AF1.  
RP PubMed=11483518; DOI=10.1093/emboj/20.15.4153;  
RX Boehm J., He Y., Greiner A., Staudt L., Wirth T.;  
RA "Regulation of BOB.1/OBF.1 stability by SIAH.";  
RL EMBO J. 20:4153-4162(2001).  
RN [15]  
RN FUNCTION IN DEGRADATION OF NUMB.  
RP PubMed=11752454; DOI=10.1073/pnas.261571998;  
RX Susini L., Passer B.J., Amzallag-Elbaz N., Juven-Gershon T., Amson R.,  
RA Prieur S., Privat N., Tudynd M., Gendron M.-C., Israeeli A., Amson R.,  
RA Oren M., Telerman A.;  
RT "Siah-1 binds and regulates the function of Numb.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:15067-15072(2001).  
RN [16]  
RN FUNCTION IN DEGRADATION OF TIEG1.  
RP PubMed=12072443; DOI=10.1074/jbc.M204812200;  
RX Johnsen S.A., Subramaniam M., Monroe D.G., Janknecht R.,  
RA Spelsberg T.C.;  
RT "Modulation of transforming growth factor beta (TGFbeta)/Smad  
RT transcriptional responses through targeted degradation of TGFbeta-  
RT inducible early gene-1 by human seven in absentia homologue.";  
RL J. Biol. Chem. 277:30754-30759(2002).  
RN [17]  
RN FUNCTION IN DEGRADATION OF SNCAIP, AND SUBCELLULAR LOCATION.  
RP PubMed=14506261; DOI=10.1074/jbc.M306347200;  
RX Negano Y., Yamashita H., Takahashi T., Kishida S., Nakamura T.,  
RA Iseki E., Hattori N., Mizuno Y., Kikuchi A., Matsumoto M.;  
RT "Siah-1 facilitates ubiquitination and degradation of synphilin-1.";  
RL J. Biol. Chem. 278:51504-51514(2003).  
RN [18]  
RN INTERACTION WITH PEG10.  
RP PubMed=12810624;  
RX Okabe H., Satoh S., Furukawa Y., Kato T., Hasegawa S., Nakajima Y.,  
RA Yamaoka Y., Nakamura Y.;  
RT "Involvement of PEG10 in human hepatocellular carcinogenesis through  
RT interaction with SIAH1.";  
RL Cancer Res. 63:3043-3048(2003).  
RN [19]  
RN TISSUE SPECIFICITY.  
RP PubMed=12557228; DOI=10.1002/gcc.10170;  
RX Matsuo K., Satoh S., Okabe H., Nomura A., Maeda T., Yamaoka Y.,  
RA Ikai I.;  
RT "SIAH1 inactivation correlates with tumor progression in  
RT hepatocellular carcinomas.";  
RL Genes Chromosomes Cancer 36:283-291(2003).  
RN [20]  
RN FUNCTION IN DEGRADATION OF RBBP8.  
RP PubMed=14654780; DOI=10.1038/sj.onc.1206994;  
RX Germani A., Prabel A., Mourah S., Podgorniak M.-P., Di Carlo A.,  
RA Ehrlich R., Gisselbrecht S., Varin-Blank N., Calvo F.,  
RA Bruzzoni-Giovanelli H.;  
RT "SIAH-1 interacts with CtIP and promotes its degradation by the  
RT proteasome pathway.";  
RL Oncogene 22:8845-8851(2003).  
RN [21]  
RN INTERACTION WITH CACBP, AND MUTANTS A; B; C; D AND E.  
RP PubMed=12421809; DOI=10.1074/jbc.M210263200;  
RX Matsuzawa S.-I., Li C., Ni C.-Z., Takayama S., Reed J.C., Ely K.R.;  
RA "Structural analysis of Siah1 and its interactions with Siah-  
RT interacting protein (SIP).";  
RL J. Biol. Chem. 278:1837-1840(2003).  
RN [22]  
RN FUNCTION IN DEGRADATION OF PML, AND MUTANTS A AND D.  
RP



```
RX PubMed=14645235; DOI=10.1074/jbc.M306407200;
RA Fanelli M., Fantozzi A., De Luca P., Caprodossi S., Matsuzawa S.-I.,
Query Match 95.1%; Score 1532; DB 1; Length 282;
Best Local Similarity 100.0%; Pred. No. 4.5e-127;
Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 MSROTATALPTGTSCPPSQRPVLPALTTASNNDLASLFCPCVCFDYLPPILQCSGHL 76
Db 1 MSROTATALPTGTSCPPSQRPVLPALTTASNNDLASLFCPCVCFDYLPPILQCSGHL 60

Qy 77 VCSNCRPKLTCCPTCRGPLGSIRNLAMEKANSVLPCKYASSCEITLPHTEKADHEEL 136
Db 61 VCSNCRPKLTCCPTCRGPLGSIRNLAMEKANSVLPCKYASSCEITLPHTEKADHEEL 120

Qy 137 CEFRPYSCPCPGASCKWQSLDAMPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 196
Db 121 CEFRPYSCPCPGASCKWQSLDAMPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 180

Qy 197 MQSCFGFHFMLVLEKQYDGHQOFFFAIVQLIGTRKQAFNAYRLELNGHRRRLTWEATP 256
Db 181 MQSCFGFHFMLVLEKQYDGHQOFFFAIVQLIGTRKQAFNAYRLELNGHRRRLTWEATP 240

Qy 257 RSHIEGATAIMNSDCLVFTSIAQLFAENGNGINVTISM 298
Db 241 RSHIEGATAIMNSDCLVFTSIAQLFAENGNGINVTISM 282

RESULT 2
O6GQJ5 PRELIMINARY; PRT; 282 AA.
ID Q6GQJ5
AC Q6GQJ5;
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC79105 protein.
GN Name=MGC79105;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyszinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
```

```
initiative."
RL Dev. Dyn. 225:384-391 (2002).
[3]
RP SEQUENCE FROM N.A.
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC072747; AAH72747.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR004162; Sina.
DR InterPro; IPR008974; Traf like.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF03145; Sina; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 282 AA; 31109 MW; 21629CB20DDFE793 CRC64;

Query Match 94.9%; Score 1529; DB 2; Length 282;
Best Local Similarity 99.3%; Pred. No. 8.3e-127;
Matches 280; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 17 MSROTATALPTGTSCPPSQRPVLPALTTASNNDLASLFCPCVCFDYLPPILQCSGHL 76
Db 1 MSROTATALPTGTSCPPSQRPVLPALTTASNNDLASLFCPCVCFDYLPPILQCSGHL 60

Qy 77 VCSNCRPKLTCCPTCRGPLGSIRNLAMEKANSVLPCKYASSCEITLPHTEKADHEEL 136
Db 61 VCSNCRPKLTCCPTCRGPLGSIRNLAMEKANSVLPCKYASSCEITLPHTEKADHEEL 120

Qy 137 CEFRPYSCPCPGASCKWQSLDAMPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 196
Db 121 CEFRPYSCPCPGASCKWQSLDAMPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 180

Qy 197 MQSCFGFHFMLVLEKQYDGHQOFFFAIVQLIGTRKQAFNAYRLELNGHRRRLTWEATP 256
Db 181 MQSCFGFHFMLVLEKQYDGHQOFFFAIVQLIGTRKQAFNAYRLELNGHRRRLTWEATP 240

Qy 257 RSHIEGATAIMNSDCLVFTSIAQLFAENGNGINVTISM 298
Db 241 RSHIEGATAIMNSDCLVFTSIAQLFAENGNGINVTISM 282

RESULT 3
SILA MOUSE
ID SIILA MOUSE STANDARD; PRT; 282 AA.
AC P61092; Q06984;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Ubiquitin ligase SIHLA (EC 6.3.2.-) (Seven in absentia homolog 1a)
DE (SiHLA) (SiHLA-1a) (mSiHLA-1a).
GN Name=SiHLA;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
RC STRAIN=Swiss; TISSUE=Eye;
RX MEDLINE=94008536; PubMed=8404535;
RA Della N.G., Senior P.V., Bowtell D.D.L.;
RT "Isolation and characterisation of murine homologues of the Drosophila
seven in absentia gene (sina).";
RL Development 117:1333-1343 (1993).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
```

complex composed of UBE2D1, SIAH1, CACYBP/SIP, SKP1A, APC and  
TbLx. Interacts with UBE2L. Interacts with alpha-tubulin.  
Interacts with PEG10, which may inhibit its activity (By  
similarity). Interacts with DAB1, which may inhibit its activity.  
SUBCELLULAR LOCATION: Cytoplasmic; predominantly. Nuclear;  
partially.  
TISSUE SPECIFICITY: Widely expressed at low level in embryos and  
adults. Expressed at higher level in testis. Due to the high  
similarity between SIAH1 and SIAH2, it is difficult to  
distinguish its own tissue specificity, suggesting that it may be  
required to modulate TP53 response. The relevance of such activity  
in vivo is however unclear and may not exist.  
DOMAIN: The RING-type zinc finger domain is essential for  
ubiquitin ligase activity.  
DOMAIN: The SBD domain (substrate-binding domain) mediates the  
homodimerization and the interaction with substrate proteins. It  
is related to the TRAF family.  
SIMILARITY: Belongs to the SINA (seven in absentia) family.  
SIMILARITY: Contains 1 C2HC-type zinc finger.  
SIMILARITY: Contains 1 RING-type zinc finger.  
SIMILARITY: Contains 1 SIAH-type zinc finger.  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; Z19579; CAA79630.1; -  
EMBL; BC046317; AAH46317.1; -  
PIR; I48763; I48763.  
PDB; 1K2F; X-ray; A/B--  
MGD; MGI:108064; Siahla.  
InterPro; IPR004162; Sina.  
InterPro; IPR008974; Traf dom.  
InterPro; IPR001841; Znf ring.  
PROSITE; PS00518; ZF\_RING\_1; FALSE\_NEG.  
PROSITE; PS00899; ZF\_RING\_2; 1.  
3D-structure; Apoptosis; Cell cycle; Ligase; Metal-binding;  
Nuclear protein; Spermatogenesis; Ubl conjugation pathway; Zinc;  
Zinc-finger.  
Zn FING 41 76 RING-type.  
ZN FING 96 121 C2HC-type.  
ZN FING 126 152 SIAH-type.  
DOMAIN 90 282 SBD.  
METAL 98 98 Zinc 1.  
METAL 105 105 Zinc 1.  
METAL 117 117 Zinc 1.  
METAL 121 121 Zinc 1.  
METAL 128 128 Zinc 2.  
METAL 135 135 Zinc 2.  
METAL 147 147 Zinc 2.  
METAL 152 152 Zinc 2.  
SEQUENCE 282 AA; 852ADCD5DD4A4FFA CRC64;  
Query Match 94.8%; Score 1528; DB 1; Length 282;  
Best Local Similarity 99.6%; Pred. No. 1e-126;  
Matches 281; Conservative 1; Mismatches 0; Gaps 0;  
QY 17 MSRTATATPTGTSKCPSPQSRVPAITGTTASNNDLASLFEPCVCFDVLPPILQCSGHL 76  
DB 1 MSRTATATPTGTSKCPSPQSRVPAITGTTASNNDLASLFEPCVCFDVLPPILQCSGHL 60  
QY 77 VCSNCRPKLTCCTCRGPLGSIRNLAMEKVANSVLPFKYVASSGCEITLPHTEKAHEEL 136  
DB 61 VCSNCRPKLTCCTCRGPLGSIRNLAMEKVANSVLPFKYVASSGCEITLPHTEKAHEEL 120  
QY 137 CEFRPSPCPFGASCKWQGSGLDAVPHLMHQHKSITTLQGEDIVFLATDINLPGADVVM 196  
DB 121 CEFRPSPCPFGASCKWQGSGLDAVPHLMHQHKSITTLQGEDIVFLATDINLPGADVVM 180

Qy 197 MQSCFGFHFMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFPAYRLELNGHRRRLTWEATP 256  
 Db 181 MQSCFGFHFMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFPAYRLELNGHRRRLTWEATP 240  
 Qy 257 RSIHEGIATAMNSDCLVFDTSIAQLFAENGLGNVTISM 298  
 Db 241 RSIHEGIATAMNSDCLVFDTSIAQLFAENGLGNVTISM 282

RESULT 4  
 SIHL\_RAT  
 ID SIHL\_RAT STANDARD; PRT; 282 AA.  
 AC Q920M5; Q06984;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Ubiquitin ligase SIAH1 (EC 6.3.2.-) (Seven in absentia homolog 1)  
 DE (Siah-1) (Siah-1a)  
 DE Name=Siah1; Synonyms=Siah1a;  
 GN Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION IN DEGRADATION OF SYP, SUBCELLULAR  
 RP LOCATION, AND INTERACTION WITH UBE2E2.  
 RC STRAIN=Sprague-Dawley;  
 RX MEDLINE=21895831; PubMed=11786535; DOI=10.1074/jbc.M107857200;  
 RA Wheeler T.C., Chin L.-S., Li Y., Roudabush F.L., Li L.;  
 RT "Regulation of synaptophysin degradation by mammalian homologues of  
 RT Seven in Absentia";  
 RL J. Biol. Chem. 277:10273-10282 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yamaguchi A., Hori O., Tohyama M.;  
 RT "Rat Siah1A";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP INTERACTIONS WITH GRM1 AND GRM5.  
 RX PubMed=10469171;  
 RA Ishikawa K., Nash S.R., Nishimune A., Neki A., Kaneko S.,  
 RA Nakanishi S.;  
 RT "Competitive interaction of seven in absentia homolog-1A and  
 RT Ca2+/calmodulin with the cytoplasmic tail of group 1 metabotropic  
 RT glutamate receptors";  
 RL Genes Cells 4:391-390 (1999).  
 CC -i- FUNCTION: E3 ubiquitin ligase protein that mediates ubiquitination  
 CC and subsequent proteasomal degradation of target proteins. E3  
 CC ubiquitin ligases accept ubiquitin from an E2 ubiquitin-  
 CC conjugating enzyme in the form of a thioester and then directly  
 CC transfers the ubiquitin to targeted substrates. Mediates E3  
 CC ubiquitin ligase activity either through direct binding to  
 CC substrates or by functioning as the essential RING domain subunit  
 CC of larger E3 complexes. Triggers the ubiquitin-mediated  
 CC degradation of many substrates, including proteins involved in  
 CC transcription regulation (MYB, POU2AF1, PML and RBBP8), a cell  
 CC surface receptor (DCC), cytoplasmic signal transduction molecules  
 CC (TRIG1 and NUBP), an antiapoptotic protein (BAG1), a microtubule  
 CC motor protein (KIF22), a protein involved in synaptic vesicle  
 CC function in neurons (SIP), a structural protein (CTNBN1) and  
 CC Ca2+/calmodulin. It is thereby involved in many cellular processes such as  
 CC apoptosis, tumor suppression, cell cycle, axon guidance,  
 CC transcription, spermatogenesis and TNF-alpha signaling. Has some  
 CC overlapping function with SIAH2.  
 CC -i- PATHWAY: Ubiquitin conjugation; third step.  
 CC -i- SUBUNIT: Homodimer. Component of some large E3 complex composed of  
 CC UBE2D1, SIAH1, CACYPBP/SIP, SKP1A, APC and TBL1X. Interacts with  
 CC UBE2I. Interacts with alpha-tubulin. Interacts with PEG10, which  
 CC may inhibit its activity (By similarity). Interacts with group 1  
 CC glutamate receptors GRM1 and GRM5. Interacts with DBA1, which may  
 CC inhibit its activity. Interacts with UBE2E2.  
 CC -i- SUBCELLULAR LOCATION: Cytoplasmic; predominantly. Nuclear;

CC partially.  
 CC -i- DOMAIN: The RING-type zinc finger domain is essential for  
 CC ubiquitin ligase activity (By similarity).  
 CC -i- DOMAIN: The SBD domain (substrate-binding domain) mediates the  
 CC homodimerization and the interaction with substrate proteins. It  
 CC is related to the TRAF family (By similarity).  
 CC -i- SIMILARITY: Belongs to the SINA (Seven in absentia) family.  
 CC -i- SIMILARITY: Contains 1 C2HC-type zinc finger.  
 CC -i- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -i- SIMILARITY: Contains 1 SIAH-type zinc finger.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF389476; AAL91362.1; -;  
 DR EMBL; AB067814; BAB70753.1; ALT\_INIT.  
 DR RGD; 620449; Siah1a.  
 DR InterPro; IPR004162; Sina.  
 DR InterPro; IPR008974; Traf\_dom.  
 DR InterPro; IPR001841; Znf\_fing.  
 DR Pfam; PF031145; Sina; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; FALSE\_NEG.  
 DR PROSITE; PS00089; ZF\_RING\_2; 1.  
 KW Apoptosis; Cell cycle; Ligase; Metal-binding; Nuclear protein;  
 KW Spermatogenesis; UbL conjugation pathway; Zinc; Zinc-finger.  
 FT ZN\_FING 41 76  
 FT ZN\_FING 96 121  
 FT ZN\_FING 126 152  
 FT ZN\_FING 152 180  
 FT DOMAIN 90 282  
 FT METAL 98 98  
 FT METAL 105 105  
 FT METAL 117 117  
 FT METAL 121 121  
 FT METAL 128 128  
 FT METAL 135 135  
 FT METAL 147 147  
 FT METAL 152 152  
 FT METAL 152 152  
 SQ SEQUENCE 282 AA; 31137 MW; 852EADC5DD44FFA CRC64;  
 Query Match 94.8%; Score 1528; DB 1; Length 282;  
 Best Local Similarity 99.6%; Pred. No. 1e-126;  
 Matches 281; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Qy 17 MSRTATALTPTGTSKCPSPSRVPAITGTTASNNDLASLFECPVCFDYVLPPILOCSGHL 76  
 Db 1 MSRTATALTPTGTSKCPSPSRVPAITGTTASNNDLASLFECPVCFDYVLPPILOCSGHL 60  
 Qy 77 VCSNCRPKLTCCCTCRGPGLSIRNLAMEKVANSVLPFCYASSGCEITLPHTEKADHEEL 136  
 Db 61 VCSNCRPKLTCCCTCRGPGLSIRNLAMEKVANSVLPFCYASSGCEITLPHTEKADHEEL 120  
 Qy 137 CEFPRYSCPCPGASCKWQSGSLDAMVPHLMHQHSITTLQGEDIVFLATDNLPGAVDWM 196  
 Db 121 CEFPRYSCPCPGASCKWQSGSLDAMVPHLMHQHSITTLQGEDIVFLATDNLPGAVDWM 180  
 Qy 197 MQSCFGFHFMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFPAYRLELNGHRRRLTWEATP 256  
 Db 181 MQSCFGFHFMLVLEKQEKYDGHQOFFAIVQLIGTRKQAEFPAYRLELNGHRRRLTWEATP 240  
 Qy 257 RSIHEGIATAMNSDCLVFDTSIAQLFAENGLGNVTISM 298  
 Db 241 RSIHEGIATAMNSDCLVFDTSIAQLFAENGLGNVTISM 282

RESULT 5  
 SIHL\_BRARE  
 ID SIHL\_BRARE STANDARD; PRT; 282 AA.  
 AC Q7ZVG6;

DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Ubiquitin ligase SIAH1 (EC 6.3.2.-) (Seven in absentia homolog 1)  
DE (Siah-1)  
GN Namesiah1;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RG NTH - Zebrafish Gene Collection (ZGC) project;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: E3 Ubiquitin ligase protein that mediates ubiquitination  
CC and subsequent proteasomal degradation of target proteins. E3  
CC ubiquitin ligases accept ubiquitin from an E2 ubiquitin-  
CC conjugating enzyme in the form of a thioester and then directly  
CC transfers the ubiquitin to targeted substrates. It probably  
CC transfers the ubiquitin-mediated degradation of different  
CC substrates (By similarity).  
CC -!- PATHWAY: Ubiquitin conjugation; third step.  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- DOMAIN: The RING-type zinc finger domain is essential for  
CC ubiquitin ligase activity (By similarity).  
CC -!- DOMAIN: The SBD domain (substrate-binding domain) mediates the  
CC homodimerization and the interaction with substrate proteins. It  
CC is related to the TRAF family (By similarity).  
CC -!- SIMILARITY: Belongs to the SINA (Seven in absentia) family.  
CC -!- SIMILARITY: Contains 1 C2HC-type zinc finger.  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
CC -!- SIMILARITY: Contains 1 SIAH-type zinc finger.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; BC045870; AAH45870.1; ALT\_INIT.  
DR HSP; BC06984; 1K2F.  
DR InterPro; IPR004162; Sina.  
DR InterPro; IPR008974; Traf\_dom.  
DR InterPro; IPR001841; Znf\_fing.  
DR Pfam; PF03145; Sina; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00518; ZF\_RING\_2; 1. FALSE\_NEG.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Ligase; Metal-binding; Ub1 conjugation pathway; Zinc; Zinc-finger.  
FT ZN FING 41 76  
FT ZN FING 96 121  
FT ZN FING 126 152  
FT ZN FING 152 182  
FT DOMAIN 90 282  
FT SBD.  
FT METAL 98 98  
FT METAL 105 105  
FT METAL 117 117  
FT METAL 121 121  
FT METAL 128 128  
FT METAL 135 135  
FT METAL 147 147  
FT METAL 152 152  
FT METAL 152 152  
SQ SEQUENCE 282 AA; 31079 MW; DC879DF8A37F58E8 CRC64;  
Query Match 93.6%; Score 1508; DB 1; Length 282;  
Best Local Similarity 97.5%; Pred. No. 5.9e-125;  
Matches 275; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
QY 17 MSROTATLPTCTSKCPSPQRVPLTGTATSNNDLASLFCPCVCFDYVLPPTLQCSGHL 76  
DB 1 MSROTATLPTGTCTSKCPSPQRVPLTGTATSNNDLASLFCPCVCFDYVLPPTLQCSGHL 60

QY 77 VCSNCRPKLTCCPTCRGPGLSIRNLAMEKVANSVLFPCKYASSGCEITLPTKADHEEL 136  
DB 61 VCSNCRPKLTCCPTCRGPGLSIRNLAMEKVANSVLFPCKYASSGCEVTLPTDKAHEEL 120  
QY 137 CEFRPYSCPCPGASCCKWGSGLDANVPHLMHQHSITTLQGEDIVFLATDINLPGAVDWM 196  
DB 121 CEFRPYSCPCPGASCCKWGSGLDANVPHLLHQHSITTLQGEDIVFLATDINLPGAVDWM 180  
QY 197 MOSCFGHFMVLVEKQEKYDGHQGFALVQLIGTRKQAFNFAIRLELNGHRRRLTWEATP 256  
DB 181 MOSCFGHFMVLVEKQEKYDGHQGFALVQLIGTRKQAFNFAIRLELNGHRRRLTWEATP 240  
QY 257 RSIHEGIATAINNSDCLVFTDSIAQLFAENGLNGLINVTISM 298  
DB 241 RSIHEGIATAINNSDCLVFTDSIAQLFAENGLNGLINVTISM 282  
RESULT 6  
SIB\_MOUSE STANDARD; PRT; 282 AA.  
AC 06985; Q7TPV6;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Ubiquitin ligase SIAH1B (EC 6.3.2.-) (Seven in absentia homolog 1b)  
DE (Siah1b) (Siah-1b).  
GN Name-Siah1b;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RC STRAIN=Swiss;  
RC MEDLINE=94008536; PubMed=8404535;  
RA Della N.G., Senior P.V., Bowtell D.D.L.;  
RT "Isolation and characterisation of murine homologues of the Drosophila  
RT seven in absentia gene (sina).";  
RL Development 117:1333-1343(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Egg;  
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP INDUCTION.

RT "Isolation of 10 differentially expressed cDNAs in p53-induced  
 RT apoptosis: activation of the vertebrate homologue of the Drosophila  
 RT seven in absentia gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:3953-3957(1996).  
 RN [4]  
 RP TISSUE SPECIFICITY.  
 RX PubMed=12842817; DOI=10.1152/ajpheart.00983.2002;  
 RA Pavre C.J., Mancuso M., Maas K., McLean J.W., Baluk P., McDonald D.M.;  
 RA "Expression of genes involved in vascular development and angiogenesis  
 RT in endothelial cells of adult lung.";  
 RT Am. J. Physiol. 285:H1917-H1938(2003).  
 RN [5]  
 RN INDUCTION.  
 RP PubMed=14985507; DOI=10.1073/pnas.0400177101;  
 RX Flucci G., Beaucourt S., Dufaut D., Lespagnol A.,  
 RA Stumpener-Cuvellette P., Gaent A., Buchwalter G., Teyndler M.,  
 RA Susini L., Lassel J.-M., Wasylyk C., Wasylyk B., Oren M., Amson R.,  
 RA Telerman A.;  
 RT "Slah-1b is a direct transcriptional target of p53: Identification of  
 RL the functional p53 responsive element in the slah-1b promoter.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:3510-3515(2004).  
 CC -1- FUNCTION: E3 Ubiquitin ligase protein that mediates ubiquitination  
 CC and subsequent proteasomal degradation of target proteins. E3  
 CC ubiquitin ligases accept ubiquitin from an E2 ubiquitin-  
 CC conjugating enzyme in the form of a thioester and then directly  
 CC transfers the ubiquitin to targeted substrates. Mediates E3  
 CC ubiquitin ligase activity either through direct binding to  
 CC substrates or by functioning as the essential RING domain subunit  
 CC of larger E3 complexes. Probably triggers the ubiquitin-mediated  
 CC degradation of many substrates (By similarity).  
 CC -1- PATHWAY: Ubiquitin conjugation; third step.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; predominantly. Nuclear;  
 CC partially (By similarity).  
 CC -1- TISSUE SPECIFICITY: Widely expressed at low level in embryos and  
 CC adults. Due to the high similarity between Slah1a and Slah1b, it  
 CC is difficult to distinguish its own tissue specificity.  
 CC Overexpressed in endothelial cells of adult lung.  
 CC -1- INDUCTION: Induced by TP53/p53, suggesting that it may be required  
 CC to modulate TP53 response.  
 CC -1- DOMAIN: The RING-type zinc finger domain is essential for  
 CC ubiquitin ligase activity.  
 CC -1- DOMAIN: The SBD domain (substrate-binding domain) mediates the  
 CC homodimerization and the interaction with substrate proteins. It  
 CC is related to the TRAF family (By similarity).  
 CC -1- SIMILARITY: Belongs to the SINA (Seven in absentia) family.  
 CC -1- SIMILARITY: Contains 1 C2HC-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -1- SIMILARITY: Contains 1 Slah-type zinc finger.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; Z19580; CAA79631.1; -;  
 CC EMBL; BC052887; AAH52887.1; -;  
 CC PIR; I48764; S35754.  
 CC HSP; Q06984; IK2F.  
 CC MGD; MG1:108063; Slah1b.  
 CC InterPro; IPR004162; Sina.  
 CC InterPro; IPR008974; Traf.dom.  
 CC InterPro; IPR001841; Znf\_ring.  
 CC Pfam; PF03145; Sina; 1.  
 CC SMART; SM00184; RING; 1.  
 CC PROSITE; PS00518; ZF\_RING\_1; FALSE\_NEG.  
 CC PROSITE; PS50089; ZF\_RING\_2; 1.  
 CC Ligase; Metal-binding; Nuclear protein; Ubi conjugation pathway; Zinc;  
 KW Zinc-finger. 41 76 RING-type.  
 FT ZN\_FING 96 121 C2HC-type.  
 FT

FT	ZN_FING	126	152	Slah-type.
FT	DOMAIN	90	282	SBD.
FT	METAL	98	98	Zinc 1 (By similarity).
FT	METAL	105	105	Zinc 1 (By similarity).
FT	METAL	117	117	Zinc 1 (By similarity).
FT	METAL	121	121	Zinc 1 (By similarity).
FT	METAL	128	128	Zinc 2 (By similarity).
FT	METAL	135	135	Zinc 2 (By similarity).
FT	METAL	147	147	Zinc 2 (By similarity).
FT	METAL	152	152	Zinc 2 (By similarity).
FT	CONFLICT	87	87	V -> M (in Ref. 1).
SQ	SEQUENCE	282 AA;	31122 MW;	A09F5D3DEEB39AC2 CRC64;

Query Match 92.7%; Score 1494; DB 1; Length 282;  
 Best Local Similarity 97.2%; Pred. No. 1e-123;  
 Matches 274; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 17 MSRTATATLPTGTSKCPSPORVPALCTGTATSNNDLASLFECPVCFDVLPPILQCSGHL 76  
 Db 1 MSQAATATLSTGTSKCPSPORVPALCTGTATSNNDLASLFECPVCFDVLPPILQCSGHL 60  
 QY 77 VCSNCRPKLTCCPTCRGLGSIIRNLAMEKVANSVLPFCVKYASSGCEITLPHTEKADHEEL 136  
 Db 61 VCSNCRPKLTCCPTCRGLGSIIRNLAMEKVANSVLPFCVKYASSGCEITLPHTEKADHEEL 120  
 QY 137 CEFRPSCPCPGASCKWQSLDAMPHLMHQHSITTTLOGEDIVFLATDINLPGAVDWM 196  
 Db 121 CEFRPSCPCPGASCKWQSLDAMPHLMHQHSITTTLOGEDIVFLATDINLPGAVDWM 180  
 QY 197 MQSCFGFHMVLVEKQKYGHOQFFAIVQLIGTRKQAEKFAYLELNGHRRRLTWEATP 256  
 Db 181 MQSCFGFHMVLVEKQKYGHOQFFAIVQLIGTRKQAEKFAYLELNGHRRRLTWEATP 240  
 QY 257 RSHIEGIATAMNSDCLVFTSTIAQLFAENGNGINVTISMC 298  
 Db 241 RSHIEGIATAMNSDCLVFTSTIAQLFAENGNGINVTISMC 282

RESULT 7  
 Q7Q5Z0 PRELIMINARY; PRT; 311 AA.  
 ID Q7Q5Z0 PRELIMINARY; PRT; 311 AA.  
 AC Q7Q5Z0; 2004 (Tremblrel. 26, Created)  
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DE EbiP5924 (Fragment).  
 GN Name=ebic5924; ORFNames=ENSANGG00000004508;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL; AAAB01008960; EAA11451.1; -;  
 DR HSP; P61092; IK2F.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0007275; P:development; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.  
 DR InterPro; IPR004162; Sina.  
 DR InterPro; IPR008974; Traf.like.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF03145; Sina; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.









CC CC homodimerization and the interaction with substrate proteins. It  
CC is related to the TRAF family (By similarity).  
CC DE -1- SIMILARITY: Belongs to the SIAH (seven in absentia) family.  
CC CC -1- SIMILARITY: Contains 1 C2HC-type zinc finger.  
CC CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
CC CC -1- SIMILARITY: Contains 1 SIAH-type zinc finger.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AB067815; BAB70754.1; -  
CC EMBL; AF389477; AAL91363.1; -  
CC HSP; Q06984; IK2F.  
CC RGD; 620778; Siah2.  
CC InterPro; IPR004162; Sina  
CC InterPro; IPR008974; Traf dom.  
CC InterPro; IPR001841; Znf\_Fing.  
CC Pfam; PF03145; Sina; 1.  
CC SMART; SM00184; RING; 1.  
CC PROSITE; PS00518; ZF\_RING\_1; FALSE\_NEG.  
CC PROSITE; PS00083; ZF\_RING\_2; 1.  
CC Apoptosis; Cell cycle; Ligase; Metal-binding; Nuclear protein;  
CC Ub1 conjugation pathway; Zinc; Zinc-finger.  
CC FT ZN\_FING 81 116 RING-type.  
CC FT ZN\_FING 137 162 C2HC-type.  
CC FT ZN\_FING 167 193 SIAH-type.  
CC FT DOMAIN 131 323 SBD.  
CC FT METAL 139 139 Zinc 1 (By similarity).  
CC FT METAL 146 146 Zinc 1 (By similarity).  
CC FT METAL 158 158 Zinc 1 (By similarity).  
CC FT METAL 162 162 Zinc 1 (By similarity).  
CC FT METAL 169 169 Zinc 2 (By similarity).  
CC FT METAL 176 176 Zinc 2 (By similarity).  
CC FT METAL 188 188 Zinc 2 (By similarity).  
CC FT METAL 193 193 Zinc 2 (By similarity).  
CC SQ SEQUENCE 325 AA; 34699 MW; 0E273AD30959982E CRC64;  
Query Match 77.4%; Score 1199; DB 1; Length 325;  
Best Local Similarity 77.6%; Pred. No. 1,4e-97;  
Matches 222; Conservative 26; Mismatches 30; Indels 8; Gaps 3;  
Qy 21 TATALTPTGTSKCPSPQV---PALTG---TASNDNLASLFECPCVDFYVLPILQCS 73  
Db 38 TISAAGFGSSAVPAAAIVISGFGAGGAGPVSPQHSLTSLFECPCVDFYVLPILQCS 97  
Qy 74 GHLVCSNCRPKLTCCPTCRGPL-GSIRNLAMEKVANSVLPCKYASSGCEITLPHTEKAD 132  
Db 98 GHLVCNCRKQLSCCPTCRGALTSPSIRNLAMEKVANSVLPCKYATGCSLTHTTEKPE 157  
Qy 133 HEELCEFRPYSCPCFGASCKWQSGSLDAMPMLHMHOKHSITTLQGEDIVFLATDINLPGAV 192  
Db 158 HEDICEYRYPSCPCFGASCKWQSGSLDAMPMLHMHOKHSITTLQGEDIVFLATDINLPGAV 217  
Qy 193 DWVMQSCFGHFHMLVLEKQEKYDCHQGFPAIVQLIGTRKQAEFPAYVELENGHRRRLTW 252  
Db 218 DWVMQSCFGHFHMLVLEKQEKYDCHQGFPAIVQLIGTRKQAEFPAYVELENGHRRRLTW 277  
Qy 253 EATPRSIHEGATATMNSDCLVFTDSIAOLFPAENGLNGLINTVITSC 298  
Db 278 EATPRSIHGDVAAALMNSDCLVFTDAIAHLFADNGLNGLINTVITSC 323  
RESULT 10  
SIA2\_HUMAN STANDARD; PRT; 324 AA.  
AC SIA255; O43270;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Ubiquitin ligase SIAH2 (EC 6.3.2.-) (seven in absentia homolog 2)  
GN (Siah-2) (hsiah2).  
GN Name-SIAH2;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1] SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RX MEDLINE=98066768; PubMed=9403064; DOI=10.1006/geno.1997.4997;  
RA Hu G., Chung T.-L., Glover T., Valentine V., Look A.T., Fearon E.R.;  
RT "Characterization of human homologs of the Drosophila seven in  
RT absentia (sina) gene.";  
RL Genomics 46:103-111(1997).  
RN [2]  
RX SEQUENCE FROM N.A., AND INTERACTION WITH VAV1.  
RX PubMed=10207103;  
RA Germani A., Romero F., Houliard M., Canonis J., Gisselbrecht S.,  
RA Fischer S., Varin-Blank N.;  
RT "hsiah2 is a new Vav binding protein which inhibits Vav-mediated  
RT signaling pathways.";  
RL Mol. Cell. Biol. 19:3798-3807(1999).  
RN [3]  
RX SEQUENCE FROM N.A.  
TX TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RX FUNCTION IN DEGRADATION OF DCC, AND INTERACTION WITH UBE2I.  
RX PubMed=9334332;  
RA Hu G., Zhang S., Vidal M., Baer J.L., Xu T., Fearon E.R.;  
RT "Mammalian homologs of seven in absentia regulate DCC via the  
RT ubiquitin-proteasome pathway.";  
RL Genes Dev. 11:2701-2714(1997).  
RN [5]  
RX FUNCTION IN DEGRADATION OF POU2AF1.  
RX PubMed=11483518; DOI=10.1093/emboj/20.15.4153;  
RA Boehm J., He Y., Greiner A., Staudt L., Wirth T.;  
RT "Regulation of BOB.1/OBF.1 stability by SIAH.";  
RL EMBO J. 20:4153-4162(2001).  
RN [6]  
RX INTERACTION WITH CACYBP.  
RX PubMed=11389839; DOI=10.1016/S1097-2765(01)00242-8;  
RA Matsuzawa S.-I., Reed J.C.;  
RT "Siah-1, SIP, and Ebi collaborate in a novel pathway for beta-catenin  
RT degradation linked to p53 responses.";  
RL Mol. Cell 7:915-926(2001).  
RN [7]  
RX FUNCTION IN DEGRADATION OF TRAF2.  
RX PubMed=12411493; DOI=10.1093/emboj/cdf576;  
RA Habelhah H., Frew I.J., Laine A., Janes P.W., Relaix F., Sassoon D.,  
RA Bowtell D.D.L., Ronai Z.;  
RT "Stress-induced decrease in TRAF2 stability is mediated by Siah2.";  
RL

EMBO J. 21:5756-5765(2002).  
 [8]  
 RN INTERACTION WITH PEG10.  
 RX PubMed:12810624;  
 RA Okabe H., Satoh S., Furukawa Y., Kato T., Hasegawa S., Nakajima Y.,  
 RA Yamaoka Y., Nakamura Y.;  
 RT "Involvement of PEG10 in human hepatocellular carcinogenesis through  
 RT interaction with SIAH1.";  
 RL Cancer Res. 63:3043-3048(2003).  
 CC -!- FUNCTION: E3 Ubiquitin ligase protein that mediates ubiquitination  
 CC and subsequent proteasomal degradation of target proteins. E3  
 CC ubiquitin ligases accept ubiquitin from an E2 ubiquitin-  
 CC conjugating enzyme in the form of a thioester and then directly  
 CC transfers the ubiquitin to targeted substrates. Mediates E3  
 CC ubiquitin ligase activity either through direct binding to  
 CC substrates or by functioning as the essential RING domain subunit  
 CC of larger E3 complexes. Triggers the ubiquitin-mediated  
 CC degradation of many substrates, including proteins involved in  
 CC transcription regulation (POU2AF1, PML, NCOR1), a cell surface  
 CC receptor (DCC), an antiapoptotic protein (BAG1), and a protein  
 CC involved in synaptic vesicle function in neurons (SYP). It is  
 CC thereby involved in apoptosis, tumor suppression, cell cycle,  
 CC transcription and signaling processes. Has some overlapping  
 CC function with SIAH1. Triggers the ubiquitin-mediated degradation  
 CC of TRAF2, whereas SIAH1 can not.  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Homodimer. Interacts with UBE2E2 (By similarity).  
 CC Interacts with VAV1, without mediating its ubiquitin-mediated  
 CC degradation. Interacts with CACYPBP/SIP. Probable component of some  
 CC large E3 complex possibly composed of UBE2D1, SIAH2, CACYBP/SIP,  
 CC SKP1A, APC and TBL1X. Interacts with PEG10, which may inhibit its  
 CC activity.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; predominantly. Nuclear;  
 CC partially (Probable).  
 CC -!- TISSUE SPECIFICITY: Widely expressed at low level.  
 CC -!- DOMAIN: The RING-type zinc finger domain is essential for  
 CC ubiquitin ligase activity.  
 CC -!- DOMAIN: The SBD domain (substrate-binding domain) mediates the  
 CC homodimerization and the interaction with substrate proteins. It  
 CC is related to the TRAF family (By similarity).  
 CC -!- SIMILARITY: Belongs to the SINA (Seven in absentia) family.  
 CC -!- SIMILARITY: Contains 1 C2HC-type zinc finger.  
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -!- SIMILARITY: Contains 1 SIAH-type zinc finger.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; Y6248; AAC51908.1; -;  
 DR EMBL; Y15268; AAT75557.1; -;  
 DR EMBL; BC013082; AAH13082.1; -;  
 DR HSSP; Q06984; 1K2F.  
 DR Genep; HGNC:10858; SIAH2.  
 DR MIM; 602213; -;  
 DR GO; GO:0005737; Cytoplasm; TAS.  
 DR GO; GO:0003714; F:transcription corepressor activity; TAS.  
 DR GO; GO:0007264; P:small GTPase mediated signal transduction; TAS.  
 DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; TAS.  
 DR InterPro; IPR004162; Sina.  
 DR InterPro; IPR008974; Traf-like.  
 DR InterPro; IPR001841; Znf\_fing.  
 DR Pfam; PF03145; Sina; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00518; 2F\_RING\_1; FALSE\_NEG.  
 DR PROSITE; PS00518; 2F\_RING\_2; 1.  
 DR Apoptosis; Cell cycle; Ligase; Metal-binding; Nuclear protein;  
 KW Ub1 conjugation pathway; Zinc; Zinc-finger.  
 FT ZN\_FING 80 115 RING-type.

FT	ZN_FING	136	161	C2HC-type.
FT	ZN_FING	166	192	SIAH-type.
FT	DOMAIN	130	322	SBD.
FT	METAL	138	138	Zinc 1 (By similarity).
FT	METAL	145	145	Zinc 1 (By similarity).
FT	METAL	157	157	Zinc 1 (By similarity).
FT	METAL	161	161	Zinc 1 (By similarity).
FT	METAL	168	168	Zinc 2 (By similarity).
FT	METAL	175	175	Zinc 2 (By similarity).
FT	METAL	187	187	Zinc 2 (By similarity).
FT	METAL	192	192	Zinc 2 (By similarity).
FT	CONFLICT	200	200	G -> E (in Ref. 1).
SO	SEQUENCE	324 AA;	34614 MW;	2D5DD845666EC924 CRC64;

Query Match 74.4%; Score 1198; DB 1; Length 324;  
 Best Local Similarity 77.3%; Pred. No. 1.7e-97;  
 Matches 221; Conservative 27; Mismatches 30; Indels 8; Gaps 2;

Qy	21	TATALPTGTSKPPSQRPVLTG-----TTASNDIASLFECPVCFDYVLPPILOQCS	73
Db	37	TISAAGPSSAVPAAAVISGPGGGGAGPVSPQHBLTSLFECPVCFDYVLPPILOQCA	96
Qy	74	GHVCSNCRPKLTCCPTCRGPL-GSIENLAMEKVANSVLPCKYASSGCEITLPHTEKAD	132
Db	97	GHVCSNCRPKLTCCPTCRGPL-TSIRNLAMEKVANSVLPCKYATTCSTLHHTKPE	156
Qy	133	HEELCEPRPSCPCGASCKWQSGSLDAVPHLMHQHKSITTLQGEDIVFLATDINLPQAV	192
Db	157	HEDICEVRPSCPCGASCKWQSGSLDAVPHLMHQHKSITTLQGEDIVFLATDINLPQAV	216
Qy	193	DWVMQSCFGFHFMLEKQEKYDGHQFPFAIVOLITRKOAEINFAYRLELNGHRRRLTW	252
Db	217	DWVMQSCFGFHFMLEKQEKYDGHQFPFAIVOLITRKOAEINFAYRLELNGHRRRLTW	276
Qy	253	EATPRSIHEGTAIWNDCLVETSTIAOLFPAENGNGINVTISMIC	298
Db	277	EATPRSIHGVAAIWNDCLVETSTIAOLFPAENGNGINVTISTC	322

RESULT 11  
 SIA2 BRARE STANDARD; PRT; 331 AA.  
 AC Q/SYJ3; Q8JH29;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Ubiquitin ligase SIAH2 (EC 6.3.2.-) (Seven in absentia homolog 2)  
 DE (Siah-2).  
 DE Name=siah2; Synonyms=siaz;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxID=7955;  
 OX [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND DEVELOPMENTAL  
 RN STAGE.  
 RC TISSUE=Embryo;  
 RX PubMed:12915316; DOI=10.1016/S1567-133X(03)00061-9;  
 RA Ro H., Kim K.E., Huh T.L., Lee S.-K., Rhee M.;  
 RT "Expression pattern of Siaz gene during the zebrafish embryonic  
 RT development.";  
 .RL Gene Expr. Patterns 3:483-488(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RN NIH - Zebrafish Gene Collection (ZGC) project;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP FUNCTION, AND DOMAIN.  
 RX PubMed:15055544;  
 RA Ro H., Jang Y., Rhee M.;  
 RT "The RING domain of Siaz, the zebrafish homologue of Drosophila seven  
 RT in absentia, is essential for cellular growth arrest.";

	Query Match	74.0%;	Score 1192.5;	DB 1;	Length 331;
	Best Local Similarity	79.5%;	Pred. No. 5.3e-97;		
	Matches 221; Conservative	23;	Mismatches 19;	Indels 15;	Gaps 3;
2y	22 ATALPTGTSKPPSQRPVALTGTASNNDLASLFCPCVFYVLPPILQCSGHLCVNSC	81		:	:
Db	68 AVALP-----VAALPG---QSLELTALFPCPVCFDYVLPPIQLCCAGHLVCNQ	113		:	:
2y	82 RPKLTCCPTCRGLP-GSIRNLAMEKVANSVLPPCKYASSGCRIETLPHTEKADHEELCEFR	140		:	:
Db	114 RQLSCOPTCRGLPTSIENLAMEKVASTLPPCKYSAGCLLSLHSEKEPEHEVCBEFR	173		:	:
2y	141 PYSCPFGASCKWQSLDAVMFLMHQHSITTLGGEDIVFLATDINLPGAVDMMQSC	200		:	:
Db	174 PYTCPFGASCKWQSGLESVMFLMHAHSITTLGGEDIVFLATDINLPGAVDMMQSC	233		:	:
2y	201 FGFHFMLVLEKBQKYDHQOFTFAIOLIGTRKOAEAFVRLNGLNGRRRLTWEATPRSIH	260		:	:
Db	234 FGFHFMLVLEKBQKYGHQOFTFAIVLLIGTRKOAEAFVRLNGLNGRRRLTWEATPRSIH	293		:	:
2y	261 EGIATAIMNSDCILVFDTSIAQLPAENGNIIGNVTISM C 298			:	:
Db	294 DGVAALIMNSDCILVFDTSIAHLFADNGLNIGNTISM C 331			:	:

RESULT 12  
SIA2A\_XENLA  
SIA2A\_XENLA STANDARD; PRT; 313 AA.  
ID AC Q918X5;  
05-JUL-2004 (Rel. 44, Created)  
05-JUL-2004 (Rel. 44, Last sequence update)  
05-JUL-2004 (Rel. 44, Last annotation update)  
Ubiquitin ligase SIAH2 (EC 6.3.2.-) (Seven in absentia homolog 2) (Xsiah-2).  
Name=XSIAH2; Synonyms=SIAH2;  
Xenopus laevis (African clawed frog).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
Xenopodinae; Xenopus.  
NCBI\_TaxID=8355;  
[1]  
SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.  
MEDLINE=21233021; PubMed=11335112; DOI=10.1016/S0925-4773(01)00332-X;  
Bogdan S., Senkel S., Esser F., Ryffel G.U., Pogge von Strandmann E.; "Misexpression of Xsiah-2 induces a small eye phenotype in Xenopus."; Mech. Dev. 103:61-69 (2001).  
-!- FUNCTION: E3 Ubiquitin ligase protein that mediates ubiquitination and subsequent proteasomal degradation of target proteins. E3 ubiquitin ligases accept ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a thioester and then directly transfers the ubiquitin to targeted substrates. Involved in eye morphogenesis, probably triggers the ubiquitin-mediated degradation of different substrates.  
-!- PATHWAY: Ubiquitin conjugation; third step.  
-!- SUBUNIT: Homodimer (By similarity).  
-!- SUBCELLULAR LOCATION: Cytoplasmic.  
-!- TISSUE SPECIFICITY: Widely expressed in early embryos until stage 40. It is then expressed in brain, spinal cord and in the developing and mature eye.  
-!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.  
-!- DOMAIN: The RING-type zinc finger domain is essential for ubiquitin ligase activity (By similarity).  
-!- DOMAIN: The SHD domain (substrate-binding domain) mediates the homodimerization and the interaction with substrate proteins. It is related to the TRAF family (By similarity).  
-!- SIMILARITY: Belongs to the SINA (seven in absentia) family.  
-!- SIMILARITY: Contains 1 C2HC-type zinc finger.  
-!- SIMILARITY: Contains 1 RING-type zinc finger.  
-!- SIMILARITY: Contains 1 SIAH-type zinc finger.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration



```
Db 37 TANTVTTGSSGCGS-----SSAGTTSSANTSSSSSSSSSSAGGDDAGMSADLTSLPFCFV 92
QY 60 CFYDVLPPILQOSGHLVCSNCRPKLTCCPTCRGPLGSTRNLAMEKVANSVLPPCKYASS 119
Db 93 CFYDVLPPILQSSGHLVCSNCRPKLTCCPTCRGPLANIRNLAMEKVANSVFPCKHSGY 152
QY 120 GCETILPHTEKADHEELCEFRPYSVPCPCGASCKWQSLDVAVPHLMHQHKSITTLQGED 179
Db 153 GCTASLVYVTEKHEETCECRPYLPCPCGASCKWQSLDVAVPHLMHQHKSITTLQGED 212
QY 180 VFLATDINLPAGVDVWVMQSCFCGHFVLEKQEKYDGHQFPFAIVOLIGTRKQOENFAY 239
Db 213 VFLATDINLPAGVDVWVMQSCFCGHFVLEKQEKYDGHQFPFAIVOLIGSRKEAENFV 272
QY 240 RLELNGHRRLLTWEATPRSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 297
Db 273 RLELNGHRRLLTWEATPRSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISL 330

RESULT 14
SINA_DROER STANDARD; PRT; 314 AA.
AC P61093;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ubiquitin ligase sina (EC 6.3.2.-) (Seven in absentia protein).
GN Name=sina;
OS Drosophila erecta (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7220;
RN [1]
RP "Assessing the impact of comparative genomic sequence data on the
RT functional annotation of the Drosophila genome."
PL Genome Biol. 3:RESEARCH0086.1-RESEARCH0086.20(2002).
CC -!- FUNCTION: E3 Ubiquitin ligase protein that is required for
CC specification of R7 photoreceptor cell fate in the eye by
CC mediating the ubiquitination and subsequent proteasomal
CC degradation of Tramtrack (ttk). E3 Ubiquitin ligases accept
CC ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a
CC thioester and then directly transfers the ubiquitin to targeted
CC substrates. Acts via the formation of a complex with ebi and phyl
CC that ubiquitinates the transcription repressor ttk, a general
CC inhibitor of photoreceptor differentiation, in a subset of
CC photoreceptor cells in the eye, leading to the differentiation of
CC cells into neurons. Also involved in external sensory organ
CC development (By similarity).
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Component of some E3 complex at least composed of sina,
CC ebi and phyl. Interacts with ebf (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- DOMAIN: The RING-type zinc finger domain is essential for
CC ubiquitin ligase activity (By similarity).
CC -!- DOMAIN: The SBD domain (substrate-binding domain) mediates the
CC interaction with substrate proteins. It is related to the TRAF
CC family (By similarity).
CC -!- SIMILARITY: Belongs to the SINA (Seven in absentia) family.
CC -!- SIMILARITY: Contains 1 C2HC-type zinc finger.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 SIAH-type zinc finger.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY190936; AAC00989.1; -.
DR FlyBase; FBgn0064599; Dere\sina.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Developmental protein; Ligase; Metal-binding; Nuclear protein;
KW Ub1 conjugation pathway; Vision; Zinc; Zinc-finger.
FT ZN FING 73 108 RING-type.
FT ZN FING 128 153 C2HC-type.
FT ZN FING 158 184 SIAH-type.
FT DOMAIN 122 314 SBD.
FT DOMAIN 14 21 Poly-Ala.
FT DOMAIN 46 55 Poly-Ser.
FT METAL 130 130 Zinc 1 (By similarity).
FT METAL 137 137 Zinc 1 (By similarity).
FT METAL 149 149 Zinc 1 (By similarity).
FT METAL 153 153 Zinc 1 (By similarity).
FT METAL 160 160 Zinc 2 (By similarity).
FT METAL 167 167 Zinc 2 (By similarity).
FT METAL 179 179 Zinc 2 (By similarity).
FT METAL 184 184 Zinc 2 (By similarity).
SQ SEQUENCE 314 AA; 33707 MW; B58D7D5E8DA2F958 CRC64;

Query Match 71.6%; Score 1153.5; DB 1; Length 314;
Best Local Similarity 75.3%; Pred. No. 1.4e-93;
Matches 217; Conservative 18; Mismatches 42; Indels 11; Gaps 1;

QY 21 TATALPTG-----TSKCPPSQRPVPAITGTTTASNNDLASLPFCPCVDFYVLPIL 69
Db 26 TWTSTGTSSAGNTSSANTSSSSSSSSAGGDDAGMSADLTSLPFCPCVDFYVLPIL 85
QY 70 QCSGHLVCSNCRPKLTCCPTCRGPLGSTRNLAMEKVANSVLPPCKYASSGCEITLPHTE 129
Db 86 QCSGHLVCSNCRPKLTCCPTCRGPLANIRNLAMEKVANSVFPCKHSGYGTASLVYTE 145
QY 130 KADHEELCEFRPYSVPCPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLP 189
Db 146 KTEHEETCECRPYLPCPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLP 205
QY 190 GAVDWMQSCFCGHFVLEKQEKYDGHQFPFAIVOLIGTRKQOENFAYRLENGHRR 249
Db 206 GAVDWMQSCFCGHFVLEKQEKYDGHQFPFAIVOLIGSRKEAENFVYRLENGRRR 265
QY 250 LTWEATPRSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 297
Db 266 LTWEATPRSHIEGATAMNSDCLVFTDTSIAQLFAENGNGINVTISL 313

RESULT 15
SINA_DROME STANDARD; PRT; 314 AA.
AC P21461; Q9VVB0;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Ubiquitin ligase sina (EC 6.3.2.-) (Seven in absentia protein).
GN Name=sina; ORFNames=CG9949;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91029488; PubMed=2146028; DOI=10.1016/0092-8674(90)90452-K;
RA Carthew R.W., Rubin G.M.;
RT "Seven in absentia, a gene required for specification of R7 cell fate
RT in the Drosophila eye.";
```

Cell 63:561-577(1990).

[2] SEQUENCE FROM N.A.

RP STRAIN=Berkley;

RC MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherzer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmei B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei G., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith H.,

RA Shue B.C., Siden-Kiamos I., Simpson M.P., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.

RT "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).

[3] GENOME REANNOTATION.

RP MEDLINE=22426069; PubMed=12537572;

RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Herman B.P.,

RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

RA Lewis S.E.

RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review."

RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).

[4] SEQUENCE FROM N.A.

RP STRAIN=Berkley; TISSUE=Head;

RC MEDLINE=22426066; PubMed=12537569;

RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,

RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,

RA Rubin G.M., Celniker S.E.

RT "A *Drosophila* full-length cDNA resource."

RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).

[5] FUNCTION IN TTK DEGRADATION, AND INTERACTIONS WITH PHYL AND EFF.

RP PubMed=9267026; DOI=10.1016/S0092-8674(00)80506-1;

RA Tang A.H., Neufeld T.P., Kwan E., Rubin G.M.

RT "PHYL acts to down-regulate TTK88, a transcriptional repressor of neuronal cell fates, by a SINA-dependent mechanism."

RL Cell 90:459-467(1997).

[6] FUNCTION IN TTK DEGRADATION, AND INTERACTION WITH PHYL.

RP PubMed=9267027; DOI=10.1016/S0092-8674(00)80507-3;

RA Li S., Li Y., Carthew R.W., Lai Z.-C.

RT "Photoreceptor cell differentiation requires regulated proteolysis of the transcriptional repressor Tramtrack."

RL Cell 90:469-478(1997).

[7] COMPONENT OF A COMPLEX WITH EBI AND PHYL.

RP PubMed=1032805; DOI=10.1093/emboj/19.20.5376;

RA Boulton S.J., Brook A., Staehling-Hampton K., Heitzler P., Dyson N.

RT "A role for Ebi in neuronal cell cycle control."

RL EMBO J. 19:5376-5386(2000).

[8] FUNCTION.

RP PubMed=11526076;

RA Pi H., Wu H.-J., Chien C.-T.

RT "A dual function of phyllopin in *Drosophila* external sensory organ development: cell fate specification of sensory organ precursor and its progeny."

RL Development 128:2699-2710(2001).

[9] FUNCTION OF THE COMPLEX.

RP PubMed=12215542; DOI=10.1128/MCB.22.19.6854-6865.2002;

RA Li S., Xu C., Carthew R.W.

RT "Phyllopin acts as an adaptor protein to link the *sina* ubiquitin ligase to the substrate protein tramtrack."

RL Mol. Cell. Biol. 22:6854-6865(2002).

CC "FUNCTION: E3 Ubiquitin ligase protein that is required for specification of R7 photoreceptor cell fate in the eye by mediating the ubiquitination and subsequent proteasomal degradation of tramtrack (ttk). E3 Ubiquitin ligases accept ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a thioester and then directly transfers the ubiquitin to targeted substrates. Acts via the formation of a complex with ebi and phyl that ubiquitinates the transcription repressor ttk, a general inhibitor of photoreceptor differentiation, in a subset of photoreceptor cells in the eye, leading to the differentiation of cells into neurons. Also involved in external sensory organ development.

CC "PATHWAY: Ubiquitin conjugation; third step.

CC "SUBUNIT: Component of some E3 complex at least composed of *sina*, ebi and phyl. Interacts with eff.

CC "SUBCELLULAR LOCATION: Nuclear and cytoplasmic.

CC "TISSUE SPECIFICITY: In many ommatidial precursor cells.

CC "DOMAIN: The RING-type zinc finger domain is essential for ubiquitin ligase activity (By similarity).

CC "DOMAIN: The SBD domain (substrate-binding domain) mediates the interaction with substrate proteins. It is related to the TRAF family (By similarity).

CC "SIMILARITY: Belongs to the SINA (seven in absentia) family.

CC "SIMILARITY: Contains 1 C2HC-type zinc finger.

CC "SIMILARITY: Contains 1 RING-type zinc finger.

CC "SIMILARITY: Contains 1 SIAH-type zinc finger.

CC "This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M38384; AAA28901.1; -

DR EMBL; AE003526; AAF49403.1; -

DR EMBL; AY060358; AAL25397.1; -

DR F01; A36195; A36195.

DR HSP; Q06984; IK2F.

DR IntAct; P21461; -

DR FlyBase; FBgn003410; *sina*.

DR GO; GO:0007423; P:sensory organ development; IMP.

DR InterPro; IPR004162; Sina.

DR InterPro; IPR008974; Traf\_like.

DR InterPro; IPR001841; Znf\_ring.

DR Pfam: PF03145; Sina: 1.  
DR PROSITE; PS00518; ZF\_RING\_1; FALSE\_NEG.  
DR PROSITE; PS0089; ZF\_RING\_2; 1.  
KW Developmental protein; Ligase; Metal-binding; Nuclear protein;  
KW Ub1 conjugation pathway; Vision; Zinc; Zinc-finger.

FT ZN\_FING 73 108 RING-type.  
FT ZN\_FING 128 153 C2HC-type.  
FT ZN\_FING 158 184 ZAH-type.  
FT DOMAIN 122 314 SBD.  
FT DOMAIN 14 21 Poly-Ala.  
FT DOMAIN 46 55 Poly-Ser.  
FT DOMAIN 26 55 Ser/Thr-rich.  
FT METAL 130 130 Zinc 1 (By similarity).  
FT METAL 137 137 Zinc 1 (By similarity).  
FT METAL 149 149 Zinc 1 (By similarity).  
FT METAL 153 153 Zinc 1 (By similarity).  
FT METAL 160 160 Zinc 2 (By similarity).  
FT METAL 167 167 Zinc 2 (By similarity).  
FT METAL 179 179 Zinc 2 (By similarity).  
FT METAL 184 184 Zinc 2 (By similarity).  
SQ SEQUENCE 314 AA; 33707 MW; B58D7D5E8A2F958 CRC64;

Query Match 71.6%; Score 1153.5; DB 1; Length 314;  
Best Local Similarity 75.3%; Pred. No. 1.4e-93;  
Matches 217; Conservative 18; Mismatches 42; Indels 11; Gaps 1;  
QY 21 TATLPTG-----TSKCPSPQRPALTTGTASNNDLASLFECPVCFDYVLPIL 69  
DB 26 TMTSTGTSSAGNTSSANTSSSSSSLSAGGAGMADLTSLECPVCFDYVLPIL 85  
QY 70 QCOSGHLVCSNCRPKLTCCPTCRGPLGSRNLAMEKVANSVLPCKYASSGCEITLPHT 129  
DB 86 QCSSGHLVCSNCRPKLTCCPTCRGPLANIRNLAMEKVANSVLPCKYSGYCTASLVY 145  
QY 130 KADHEELCEFRPVSQPCGASCKWQGSIDAVPHLMHQHSITLQGEDIVFLATDINLP 189  
DB 146 KTEHEETCECRPYLCPCGASCKWQGSIDAVPHLMHQHSITLQGEDIVFLATDINLP 205  
QY 190 GAVDWMQSCFGHFMLVLEKQYDGHQOPFAIVOLIGTRKOAENFAYRLELNGHRR 249  
DB 206 GAVDWMQSCFGHFMLVLEKQYDGHQOPFAIVOLIGTRKOAENFAYRLELNGHRR 265  
QY 250 LTWEATPRSIHEGIATAIMNSDCLVFTSIAQLFAENGLGINVTISM 297  
DB 266 LTWEATPRSIHEGIVASAIHNSDCLVFTSIAQLFAENGLGINVTISL 313

Search completed: April 25, 2005, 06:38:42  
Job time : 85 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 05:57:40 ; Search time 25 Seconds  
(without alignments)  
889.817 Million cell updates/sec

Title: US-10-679-246-2  
Perfect score: 1611  
Sequence: 1 MWIIIFLLPPYVFIEMSRQ.....IAQLFAENGNGINVTISM 298

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUTS-COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	298	US-09-591-694-2	Sequence 2, Appli
2	1516	94.1	282	US-09-544-618-17	Sequence 17, Appli
3	1512	93.9	282	US-09-544-618-18	Sequence 18, Appli
4	1489	92.4	282	US-09-544-618-12	Sequence 12, Appli
5	1482	92.0	282	US-09-544-618-19	Sequence 19, Appli
6	1122.5	69.7	314	US-09-544-618-20	Sequence 20, Appli
7	447	27.7	313	US-09-362-506-2	Sequence 2, Appli
8	326	20.2	325	US-09-325-932A-83	Sequence 83, Appli
9	325	20.2	325	US-09-325-932A-201	Sequence 201, Appli
10	123.5	7.7	299	US-09-270-767-45308	Sequence 45308, A
11	121.5	7.5	172	US-09-270-767-60801	Sequence 60801, A
12	107.5	6.7	120	US-09-325-932A-84	Sequence 84, Appli
13	107	6.6	485	US-08-724-394A-8	Sequence 8, Appli
14	106	6.6	522	US-08-639-237-2	Sequence 2, Appli
15	106	6.6	522	US-08-375-405-2	Sequence 2, Appli
16	105.5	6.5	22	US-08-012-269A-4	Sequence 4, Appli
17	102.5	6.4	474	US-09-949-016-7120	Sequence 7120, Ap
18	99.5	6.2	298	US-09-127-928-2	Sequence 2, Appli
19	99.5	6.2	299	US-09-949-016-7820	Sequence 7820, Ap
20	97.5	6.1	50	US-09-270-767-34977	Sequence 34977, A
21	97.5	6.1	50	US-09-270-767-50194	Sequence 50194, A
22	96.5	6.0	326	US-09-538-092-1019	Sequence 1019, Ap
23	96.5	6.0	343	US-09-949-016-11007	Sequence 11007, A
24	96	6.0	913	US-09-248-796A-15734	Sequence 15734, A
25	93.5	5.8	366	US-09-908-988B-2	Sequence 2, Appli
26	93.5	5.8	487	US-09-949-016-11205	Sequence 11205, A
27	93.5	5.8	600	US-09-327-984A-6	Sequence 6, Appli

28	92	5.7	600	3	US-09-212-971-12	Sequence 12, Appli
29	92	5.7	600	3	US-08-800-929A-12	Sequence 12, Appli
30	92	5.7	600	3	US-09-617-053A-12	Sequence 12, Appli
31	91.5	5.7	447	4	US-09-949-016-7884	Sequence 7884, Ap
32	91.5	5.7	545	4	US-09-908-988B-4	Sequence 4, Appli
33	91.5	5.7	664	4	US-09-949-016-9437	Sequence 9437, Ap
34	90	5.6	539	4	US-09-949-016-6363	Sequence 6363, Ap
35	90	5.6	602	4	US-09-201-936-40	Sequence 40, Appli
36	90	5.6	602	4	US-09-011-356-40	Sequence 40, Appli
37	90	5.6	602	4	US-09-672-717-227	Sequence 227, App
38	90	5.6	602	4	US-09-201-932-40	Sequence 40, Appli
39	89.5	5.6	475	4	US-09-949-016-6317	Sequence 6317, Ap
40	89	5.5	33	3	US-09-324-455-18	Sequence 18, Appli
41	89	5.5	327	3	US-09-949-016-7575	Sequence 7575, Ap
42	89	5.5	383	3	US-08-857-076-105	Sequence 105, App
43	89	5.5	568	4	US-09-949-016-6339	Sequence 6339, Ap
44	89	5.5	568	4	US-08-813-323C-2	Sequence 2, Appli
45	89	5.5	1964	3	US-09-467-997-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-591-694-2  
; Sequence 2, Application US/09591694  
; Patent No. 6638734  
; GENERAL INFORMATION:  
; APPLICANT: John C. Reed  
; APPLICANT: Shu-ichi Matsuzawa  
; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved  
; TITLE OF INVENTION: in Protein Degradation, Products and Methods Related Thereto  
; FILE REFERENCE: P-LJ 4220  
; CURRENT APPLICATION NUMBER: US/09/591,694  
; CURRENT FILING DATE: 2000-06-09  
; EARLIER APPLICATION NUMBER: US 09/330,517  
; EARLIER FILING DATE: 1999-06-11  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-591-694-2

Query Match	100.0%;	Score 1611;	DB 4;	Length 298;
Best Local Similarity	100.0%;	Pred. No. 3.2e-159;		
Matches 298;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MWIIIFLLPPYVFIEMSRQTATALTGTSCPPSQRVPALTGTTASNNDLASLFCPCVC	60	
DB	1	MWIIIFLLPPYVFIEMSRQTATALTGTSCPPSQRVPALTGTTASNNDLASLFCPCVC	60	
QY	61	PDYVLPPILOCSGHLVCSNCPKLTCCPTCGPLGSIRNLAMEKVANSVLPCKYASSG	120	
DB	61	PDYVLPPILOCSGHLVCSNCPKLTCCPTCGPLGSIRNLAMEKVANSVLPCKYASSG	120	
QY	121	CSITIPHTKADHEELCEPRYSCPCPGASCCKQGSGLDAMPHLMHOKHSITTLQGEDIV	180	
DB	121	CSITIPHTKADHEELCEPRYSCPCPGASCCKQGSGLDAMPHLMHOKHSITTLQGEDIV	180	
QY	181	FLATDINLPGADVMMWQSCFGFHEMLVLEKEKYDGHQFFAIVQLIGTRKQAFNAYR	240	
DB	181	FLATDINLPGADVMMWQSCFGFHEMLVLEKEKYDGHQFFAIVQLIGTRKQAFNAYR	240	
QY	241	LELNGHRRRLTWEATPRSIHEGIATAIMNSDCLVFDTSIAQLFAENGNGINVTISM	298	
DB	241	LELNGHRRRLTWEATPRSIHEGIATAIMNSDCLVFDTSIAQLFAENGNGINVTISM	298	

RESULT 2  
US-09-544-618-17  
; Sequence 17, Application US/09544618

```

; Patent No. 6503502
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Cohen, Daniel
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, PROTEINS, DRUGS AND DIAGNOSTIC
; TITLE OF INVENTION: AGENTS OF USE IN TREATING CANCER
; FILE REFERENCE: 065691-0139
; CURRENT APPLICATION NUMBER: US/09/544,618
; CURRENT FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 17
; LENGTH: 282
; TYPE: PRT
; ORGANISM: HUMAN
US-09-544-618-17

Query Match 94.1%; Score 1516; DB 4; Length 282;
Best Local Similarity 99.3%; Pred. No. 2.2e-149;
Matches 280; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 76
DB 1 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 60
QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136
DB 61 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120
QY 137 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 196
DB 121 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 180
QY 197 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 256
DB 181 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 240
QY 257 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 298
DB 241 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 282
; TYPE: PRT
; ORGANISM: MOUSE
US-09-544-618-18

Query Match 93.9%; Score 1512; DB 4; Length 282;
Best Local Similarity 98.9%; Pred. No. 5.8e-149;
Matches 279; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 17 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 76
DB 1 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 60
QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136
; TYPE: PRT
; ORGANISM: MOUSE
US-09-544-618-19

Query Match 92.4%; Score 1489; DB 4; Length 282;
Best Local Similarity 96.8%; Pred. No. 1.4e-146;
Matches 273; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 17 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 76
DB 1 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 60
QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136
DB 61 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120
QY 137 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 196
DB 121 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 180
QY 197 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 256
DB 181 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 240
QY 257 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 298
DB 241 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 282
; TYPE: PRT
; ORGANISM: TSAP 3
US-09-544-618-12

Query Match 92.4%; Score 1489; DB 4; Length 282;
Best Local Similarity 96.8%; Pred. No. 1.4e-146;
Matches 273; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 17 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 76
DB 1 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 60
QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136
DB 61 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120
QY 137 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 196
DB 121 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 180
QY 197 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 256
DB 181 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 240
QY 257 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 298
DB 241 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 282
; TYPE: PRT
; ORGANISM: TSAP 3
US-09-544-618-12
```

```

; Patent No. 6503502
; GENERAL INFORMATION:
; APPLICANT: Telerman, Adam
; APPLICANT: Amson, Robert
; APPLICANT: Cohen, Daniel
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, PROTEINS, DRUGS AND DIAGNOSTIC
; TITLE OF INVENTION: AGENTS OF USE IN TREATING CANCER
; FILE REFERENCE: 065691-0139
; CURRENT APPLICATION NUMBER: US/09/544,618
; CURRENT FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 12
; LENGTH: 282
; TYPE: PRT
; ORGANISM: TSAP 3
US-09-544-618-12

Query Match 92.4%; Score 1489; DB 4; Length 282;
Best Local Similarity 96.8%; Pred. No. 1.4e-146;
Matches 273; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 17 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 76
DB 1 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 60
QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136
DB 61 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120
QY 137 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 196
DB 121 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 180
QY 197 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 256
DB 181 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 240
QY 257 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 298
DB 241 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 282
; TYPE: PRT
; ORGANISM: TSAP 3
US-09-544-618-12

Query Match 92.4%; Score 1489; DB 4; Length 282;
Best Local Similarity 96.8%; Pred. No. 1.4e-146;
Matches 273; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 17 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 76
DB 1 MSRTATATLPTGTSKCPSPQSRVPALTGTASNNDLASLFECPVCFDYVLPPILOQCSGHL 60
QY 77 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136
DB 61 VCSNCRPKLTCCPTCRGLSGIRNLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120
QY 137 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 196
DB 121 CEFRPYSCEPCGASCKWQSLDVAVPHLMHQHKSITTLQGEDIVFLATDINLPGAVDWM 180
QY 197 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 256
DB 181 MOSCFGHFMVLVEKQEKYDGHQOFFAIVOLIGTRKQAFNFAIRLELNGHRRRLTWEATP 240
QY 257 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 298
DB 241 RSHIEGIATAIMNSDCLVFDTSIAQLFAENGLGINVTISM 282
; TYPE: PRT
; ORGANISM: TSAP 3
US-09-544-618-12
```

;; CURRENT APPLICATION NUMBER: US/09/544,618  
;; CURRENT FILING DATE: 2000-03-06  
;; NUMBER OF SEQ ID NOS: 22  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 19  
;; LENGTH: 282  
;; TYPE: PRT  
;; ORGANISM: MOUSE  
US-09-544-618-19

Query Match 92.0%; Score 1482.; DB 4; Length 282;  
Best Local Similarity 96.8%; Pred. No. 7.7e-146;  
Matches 273; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 17 MSRTATALPTGTSKPPSORVPALTGTTASNNDLASLFECPVCFDYVLPPIQCSGHL 76  
DB 1 MSRQAATLSTGTSKPPSORVPALTDTTASNNDLASLFECPVCFDYVLPPIQCSGHL 60  
QY 77 VCSNCRPKLTCCPTCRGPIGSRINLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 136  
DB 61 VCSNCRPKLTCCPTCRGPIGSRINLAMEKVANSVLPCKYASSGCEITLPHTEKADHEEL 120  
QY 137 CEFRPYSCPCGASCKWQSLDAMPHLMHOKSIITTLQGEDIVFLATDINLPGVDWVM 196  
DB 121 CEFRPYSCPCGASCKWQSLDAMPHLMHOKSIITTLQGEDIVFLATDINLPGVDWVM 180  
QY 197 MQSCFGFHFMLVLEKQEKYDGHQFFAIVQLIGTRKQAFNFAIRLELNGHRRRLTWEATP 256  
DB 181 MQSCFGFHFMLVLEKQEKYDGHQFFAIVQLIGTRKQAFNFAIRLELNGHRRRLTWEATP 240  
QY 257 RSIHEGIATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 298  
DB 241 RSIHEGIATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 282

## RESULT 6

US-09-544-618-20  
;; Sequence 20, Application US/09544618  
;; Patent No. 6503502  
;; GENERAL INFORMATION:  
;; APPLICANT: Telerman, Adam  
;; APPLICANT: Anson, Robert  
;; APPLICANT: Cohen, Daniel  
;; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, PROTEINS, DRUGS AND DIAGNOSTIC  
;; FILE REFERENCE: AGENTS OF USE IN TREATING CANCER  
;; FILE REFERENCE: 065691-0139  
;; CURRENT APPLICATION NUMBER: US/09/544,618  
;; CURRENT FILING DATE: 2000-03-06  
;; NUMBER OF SEQ ID NOS: 22  
;; SOFTWARE: Patent In Ver. 2.0  
;; SEQ ID NO 20  
;; LENGTH: 314  
;; TYPE: PRT  
;; ORGANISM: DROSINA  
US-09-544-618-20

Query Match 69.7%; Score 1122.5; DB 4; Length 314;  
Best Local Similarity 73.0%; Pred. No. 2.3e-108;  
Matches 211; Conservative 21; Mismatches 44; Indels 13; Gaps 1;  
QY 22 ATALPTGTSKPPSORVPALTGTTASNN-----DLASLFECPVCFDYVLPPI 68  
DB 25 ATNTSTGSSAGNTSSAMTSSSSSSLSAGGGGAGMSADLTSLFECPVCFDYVLPPI 84  
QY 69 LQCSGHLVCSNCRPKLTCCPTCRGPIGSRINLAMEKVANSVLPCKYASSGCEITLPH 128  
DB 85 LQCSGHLVCSNCRPKLTCCPTCRGPIGSRINLAMEKVANSVLPCKYASSGCEITLPH 144  
QY 129 EKADHEELCEFRPYSCPCGASCKWQSLDAMPHLMHOKSIITTLQGEDIVFLATDINL 188  
DB 145 EKTEHEETCEFRPYSCPCGASCKWQSLDAMPHLMHOKSIITTLQGEDIVFLATDINL 204  
QY 189 PGAVDWMQSCFGFHFMLVLEKQEKYDGHQFFAIVQLIGTRKQAFNFAIRLELNGHRR 248

DB 205 PGAVDWMQSCFGFHFMLVLEKQEKYDGHQFFAIVQLIGTRKQAFNFAIRLELNGHRR 264  
QY 249 RLTTWEATPRSIHEGIATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 297  
DB 265 RLTTWEATPRSIHEGIATAMNSDCLVFTDTSIAQLFAENGNGINVTISM 313  
RESULT 7  
US-09-362-506-2  
;; Sequence 2, Application US/09362506  
;; Patent No. 6111167  
;; GENERAL INFORMATION:  
;; APPLICANT: Mahajan, Pramod B.  
;; TITLE OF INVENTION: Maize SINA Orthologue-1 and Uses Thereof  
;; FILE REFERENCE: 0936  
;; CURRENT APPLICATION NUMBER: US/09/362,506  
;; CURRENT FILING DATE: 1999-07-27  
;; EARLIER APPLICATION NUMBER: 60/100,258  
;; EARLIER FILING DATE: 1998-09-14  
;; NUMBER OF SEQ ID NOS: 3  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 2  
;; LENGTH: 313  
;; TYPE: PRT  
;; ORGANISM: Zea mays  
US-09-362-506-2

Query Match 27.7%; Score 447; DB 3; Length 313;  
Best Local Similarity 37.6%; Pred. No. 4.3e-38;  
Matches 103; Conservative 44; Mismatches 101; Indels 26; Gaps 9;  
QY 39 PALTGTTASNNDLAS-----LFECPCVCFDYVLPPIQCSGHLVCSNCRPKL-TCCPTC 91  
DB 39 PASTSASLANVALSSLCGLNDLLECPVCTNSMRPPILOCPNGHTICSSCKHRVENHCPTC 98  
QY 92 RGPLGSIRINLAMEKVANSVLPCKYASSGCEITLPHTEKADHEELCEFRPYSCPCGASC 151  
DB 99 ROELGNIRCLAEKVAEQQLFECKYQSTGCTRIHPYKSKLKHHEELCFRFPYSCPYAGSEC 158  
QY 152 KWQGSILDAVPHLMHOKSIITTLQGEDIVFLATDINLPGVD---WVM-MQSCFGFHFML 207  
DB 159 LIAGDVPLVLSLINDHK-VDLHEGCTFNHRVTKPM-PYEVENATWMLTVFKCFGQHFL 216  
QY 208 -----VLEKQEKYDGHQFFAIVQLIGTRKQAFNFAIRLELNGHRRRLTWEATPRSIHEG 262  
DB 217 HFEAFVLGMAVY-----MAFLRFMGESSEAQFGYSLVGGGGRKLTWQGTFRSVRDS 270  
QY 263 IATAMNSDCLVFTDTSIAQLFAENG--NLGINVT 294  
DB 271 HRKVRDSFDGLIHRNMALFFSGGGRQELKLRVT 304

## RESULT 8

US-09-325-932A-83  
;; Sequence 83, Application US/09325932A  
;; Patent No. 6451604  
;; GENERAL INFORMATION:  
;; APPLICANT: Flinn, Barry  
;; APPLICANT: Lasham, Annette  
;; TITLE OF INVENTION: Compositions affecting programmed cell  
;; FILE REFERENCE: death and their use in the modification of forestry plant develop  
;; FILE REFERENCE: 1022  
;; CURRENT APPLICATION NUMBER: US/09/325,932A  
;; CURRENT FILING DATE: 1999-06-04  
;; NUMBER OF SEQ ID NOS: 206  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 83  
;; LENGTH: 325  
;; TYPE: PRT  
;; ORGANISM: Pinus radiata  
US-09-325-932A-83

Query Match	7.5%;	Score 121.5;	DB 4;	Length 172;
Best Local Similarity	22.4%;	Pred. No. 0.00013;		
Matches	41;	Conservative 22;	Mismatches 55;	Indels 65; Gaps 8
QY	63	YVLPILQCGSHLVCSNCRKPLTCCPTCRGPLG--SIRNLAMEKVANSLVFPCKYASS	119	
DB	16	YSLPSL-----RDIATPCNCRVEISKSTASRNLAWEKAELSPSEQCF	59	
QY	120	GCETTLPHTEKADHEE-LCEPRPYSCPCFGASCWQSGSLDAVMPH--LMHOHKSITLLQ	175	
DB	60	-CNKEFPYKSLRHQHQCQRPYCKYHRIQCQWRGYPHETNEHERNCLHPQKS----	113	
QY	176	GEDIVFLATDINLPGADVMMWQSGFGFHFMLVLE-----KQEKYDGHQGFALVQLIG	229	



NAME: Brezner, David J  
REGISTRATION NUMBER: 24,7747  
REFERENCE/DOCKET NUMBER: T96-004/A63312  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 522 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-975-405-2

Query Match 6.6%; Score 106; DB 1; Length 522;  
Best Local Similarity 19.3%; Pred. No. 0.027;  
Matches 72; Conservative 43; Mismatches 115; Indels 144; Gaps 17;

QY 30 SKCPPSQRPVPAITGTTASNN-----DLALFECPCVCFDYVLPPILOQ 71  
DB 25 SSCSAVTKDDSVGGTASTGNLSSSFMEETIQGYDVEFDPPLESKEYECPICL-MALREAVQT 83  
QY 72 QSGHLVCSNCRPKLTCCPTCRGPLGSIRN-----LAMEKVANSVLPCKYA----- 117  
DB 84 PCGRHFCACIILK-----SIRDAHGKCPVDNEILLENQLFPDNFAKREILSLM 131  
QY 118 ----SSGC--EITLPHTEKADHEELCEFRPYSCP-----CPGASCKW 153  
DB 132 VKCPNEGCLHMKWELRHLE--DQAHCEPALMDCPQCQRPFOKFHINIHLKDCP---RR 185  
QY 154 QGSLDAMPHLMHQHSITITLOGEDIVFLATDINLPGA-----VDWVMQ----- 198  
DB 186 QVSCDNCNCAASMAFEDKEI-----HDQNCPLANVCEYCNLTILIREQMPNHYDL 233  
QY 199 -----SCFGFHEM-----LVLEKQEKYDGHQOQFFA-----IVOLIGTRKQAE 235  
DB 234 DCPATPICTFTFCHEKQKQNRHLARHLQENTQSHRMMLAQAVHLSVIPDSGYISEVR 293  
QY 236 NFAY-----RLELNGHRRRLTWEATPSIHGATATAMNSDCLVFDTSIAQLFAEN-- 286  
DB 294 NFQETIHQLEGLRVRQDQHIR---ELTAKMETQSMYVSELKRTITRTLEDKVAIEAQOQN 350  
QY 287 -----GNLGINV 293  
DB 351 GIYIWKIGNFGMHL 364

Search completed: April 25, 2005, 06:39:44  
Job time : 28 secs

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 522 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-639-237-2

Query Match 6.6%; Score 106; DB 1; Length 522;  
Best Local Similarity 19.3%; Pred. No. 0.027;  
Matches 72; Conservative 43; Mismatches 115; Indels 144; Gaps 17;

QY 30 SKCPPSQRPVPAITGTTASNN-----DLALFECPCVCFDYVLPPILOQ 71  
DB 25 SSCSAVTKDDSVGGTASTGNLSSSFMEETIQGYDVEFDPPLESKEYECPICL-MALREAVQT 83  
QY 72 QSGHLVCSNCRPKLTCCPTCRGPLGSIRN-----LAMEKVANSVLPCKYA----- 117  
DB 84 PCGRHFCACIILK-----SIRDAHGKCPVDNEILLENQLFPDNFAKREILSLM 131  
QY 118 ----SSGC--EITLPHTEKADHEELCEFRPYSCP-----CPGASCKW 153  
DB 132 VKCPNEGCLHMKWELRHLE--DQAHCEPALMDCPQCQRPFOKFHINIHLKDCP---RR 185  
QY 154 QGSLDAMPHLMHQHSITITLOGEDIVFLATDINLPGA-----VDWVMQ----- 198  
DB 186 QVSCDNCNCAASMAFEDKEI-----HDQNCPLANVCEYCNLTILIREQMPNHYDL 233  
QY 199 -----SCFGFHEM-----LVLEKQEKYDGHQOQFFA-----IVOLIGTRKQAE 235  
DB 234 DCPATPICTFTFCHEKQKQNRHLARHLQENTQSHRMMLAQAVHLSVIPDSGYISEVR 293  
QY 236 NFAY-----RLELNGHRRRLTWEATPSIHGATATAMNSDCLVFDTSIAQLFAEN-- 286  
DB 294 NFQETIHQLEGLRVRQDQHIR---ELTAKMETQSMYVSELKRTITRTLEDKVAIEAQOQN 350  
QY 287 -----GNLGINV 293  
DB 351 GIYIWKIGNFGMHL 364

RESULT 15  
US-08-975-405-2  
Sequence 2, Application US/08975405  
Patent No. 5767244  
GENERAL INFORMATION:  
APPLICANT: Goeddel, David V.  
APPLICANT: Xiong, Jessie  
TITLE OF INVENTION: No. 5767244el Protein - TRAF6  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/975,405  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/639,237  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 25, 2005, 06:38:49 ; Search time 403 Seconds  
(without alignments)  
246.084 Million cell updates/sec

Title: US-10-679-246-2

Perfect score: 1611

Sequence: 1 MWIIIFLLPPYVFISEMSRQ.....IAQLFAENGLNINTWISMC 298

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	298	16	US-10-679-246-2
2	1603	99.5	298	15	US-10-108-260A-2712
3	1198	74.4	324	14	US-10-177-293-425
4	1190	73.9	324	16	US-10-755-889-124
5	462.5	28.7	308	16	US-10-437-963-185899
6	432	26.8	89	9	US-09-925-297-543
7	422	26.2	402	15	US-10-437-963-187446
8	419	26.0	322	16	US-10-424-599-241141
9	417	25.9	305	15	US-10-425-114-43471
10	416	25.8	324	15	US-10-424-599-164340
11	415	25.8	693	16	US-10-437-963-111838
12	413.5	25.7	311	15	US-10-424-599-161695
13	411	25.5	309	16	US-10-324-120-1

14	407.5	25.3	498	16	US-10-437-963-158317	Sequence 158317,
15	405	25.1	309	16	US-10-324-120-3	Sequence 3, Appli
16	404	25.1	309	15	US-10-424-599-158683	Sequence 158683,
17	401	24.9	309	16	US-10-324-120-2	Sequence 2, Appli
18	398.5	24.7	375	16	US-10-437-963-132004	Sequence 132004,
19	397	24.6	310	15	US-10-424-599-161694	Sequence 161694,
20	392	24.3	327	15	US-10-424-599-164338	Sequence 164338,
21	332	20.6	192	16	US-10-767-701-34707	Sequence 34707, A
22	326	20.2	325	14	US-10-219-220-83	Sequence 83, Appl
23	326	20.2	325	14	US-10-219-220-83	Sequence 270, App
24	325	20.2	325	14	US-10-219-220-201	Sequence 201, App
25	307.5	19.1	184	15	US-10-108-260A-3779	Sequence 3779, Ap
26	276.5	17.2	187	16	US-10-767-701-50920	Sequence 50920, A
27	260.5	16.2	223	16	US-10-437-963-127636	Sequence 127636,
28	237.5	14.7	358	15	US-10-424-599-271409	Sequence 271409,
29	236.5	14.7	214	16	US-10-767-701-41512	Sequence 41512, A
30	235.5	14.6	341	15	US-10-425-114-67688	Sequence 67688, A
31	235.5	14.6	417	16	US-10-437-963-169168	Sequence 169168,
32	227	14.1	304	16	US-10-437-963-180489	Sequence 180489,
33	225	14.0	292	16	US-10-437-963-172205	Sequence 172205,
34	211.5	13.1	151	16	US-10-767-701-53418	Sequence 53418, A
35	203.5	12.6	307	16	US-10-437-963-173045	Sequence 173045,
36	178	11.0	159	16	US-10-767-701-38997	Sequence 38997, A
37	174.5	10.8	128	16	US-10-437-963-181637	Sequence 181637,
38	173.5	10.8	285	15	US-10-425-114-65663	Sequence 65663, A
39	170.5	10.6	203	16	US-10-437-963-113170	Sequence 113170,
40	169.5	10.5	90	15	US-10-424-599-273575	Sequence 273575,
41	169.5	10.5	332	16	US-10-437-963-131170	Sequence 131170,
42	168	10.4	261	16	US-10-437-963-159080	Sequence 159080,
43	165.5	10.3	386	16	US-10-437-963-194113	Sequence 194113,
44	159	9.9	289	16	US-10-437-963-179909	Sequence 179909,
45	154.5	9.6	105	16	US-10-767-701-60980	Sequence 60980, A

ALIGNMENTS

RESULT 1

US-10-679-246-2  
; Sequence 2, Application US/10679246  
; Publication No. US20040163138A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved  
; TITLE OF INVENTION: in Protein Degradation, Products and Methods Related Thereto  
; FILE REFERENCE: 66821-235  
; CURRENT APPLICATION NUMBER: US/10/679,246  
; PRIOR FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: US 09/591,694  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-679-246-2

Query Match	100.0%	Score 1611;	DB 16;	Length 298;
Best Local Similarity	100.0%	Pred. No. 2.3e-148;		
Matches 298;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MVIIIFLLPPYVFISEMSRQTATLPTGTSKCPSPORVPALGTGTASNNDLASLRECPVC	60	
Db	1	MVIIIFLLPPYVFISEMSRQTATLPTGTSKCPSPORVPALGTGTASNNDLASLRECPVC	60	
QY	61	FDYVLPPIIQCOSGHLVCSNCRPKLTCCPTCRGLGSIIRNLAMEKVANSVLPFCYASSG	120	
Db	61	FDYVLPPIIQCOSGHLVCSNCRPKLTCCPTCRGLGSIIRNLAMEKVANSVLPFCYASSG	120	
QY	121	CSITLPHTEKADHELCFRPVCPCPGASCWKQOSLDAMPHLMHQHSITTLQGEDIV	180	



Db 121 CEITLPHTEKADHEELCEPRPSCPCPGASCKWQSLDAMVPHLMHQHKSITTLQGEDIV 180  
Qy 181 FLATDINLPGADVMMQSCFGFHFMLVLEKQKDYDGHQOFAIVOLIGTRKQAEAFAYR 240  
Db 181 FLATDINLPGADVMMQSCFGFHFMLVLEKQKDYDGHQOFAIVOLIGTRKQAEAFAYR 240  
Qy 241 LEELNGHRRRLTWEATPRSHIEGIATAIMNSDCLVFDTSIAQLFAENGNGINVTISM 298  
Db 241 LEELNGHRRRLTWEATPRSHIEGIATAIMNSDCLVFDTSIAQLFAENGNGINVTISM 298

## RESULT 2

US-10-108-260A-2712  
; Sequence 2712, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2712  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-2712

Query Match 99.5%; Score 1603; DB 15; Length 298;  
Best Local Similarity 99.7%; Pred. No. 1.4e-147;  
Matches 297; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MWIIIFLLPPYFISEMSQTATLPTGTSKPPSORVPALTGTASNNDLASLSECPVC 60  
Db 1 MWIIIFLLPPYFISEMSQTATLPTGTSKPPSORVPALTGTASNNDLASLSECPVC 60  
Qy 61 FDYVLPPILOQSGHLVCSNCRPKLTCCPTCRGLGSIIRNLAMEKVANSVLPFCYASSG 120  
Db 61 FDYVLPPILOQSGHLVCSNCRPKLTCCPTCRGLGSIIRNLAMEKVANSVLPFCYASSG 120  
Qy 121 CEITLPHTEKADHEELCEPRPSCPCPGASCKWQSLDAMVPHLMHQHKSITTLQGEDIV 180  
Db 121 CEITLPHTEKADHEELCEPRPSCPCPGASCKWQSLDAMVPHLMHQHKSITTLQGEDIV 180  
Qy 181 FLATDINLPGADVMMQSCFGFHFMLVLEKQKDYDGHQOFAIVOLIGTRKQAEAFAYR 240  
Db 181 FLATDINLPGADVMMQSCFGFHFMLVLEKQKDYDGHQOFAIVOLIGTRKQAEAFAYR 240  
Qy 241 LEELNGHRRRLTWEATPRSHIEGIATAIMNSDCLVFDTSIAQLFAENGNGINVTISM 298  
Db 241 LEELNGHRRRLTWEATPRSHIEGIATAIMNSDCLVFDTSIAQLFAENGNGINVTISM 298

## RESULT 3

US-10-177-293-425  
; Sequence 425, Application US/10177293  
; Publication No. US20030124128A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Gannavarpu, Manjula  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Mertens, Maureen  
; APPLICANT: Myer, Vic  
; APPLICANT: Wang, Yongzhen  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Monahan, John  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Bast Jr., Robert C.  
; APPLICANT: Hortobagyi, Gabriel N.

; APPLICANT: Pusztai, Lajos  
; APPLICANT: Mexic, Funda  
; APPLICANT: Sahin, Aysegul  
; APPLICANT: Mills, Gordon B.  
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,  
; FILE REFERENCE: PREVENTION, AND THERAPY OF BREAST CANCER  
; TITLE OF INVENTION: MRI-038  
; CURRENT APPLICATION NUMBER: US/10/177,293  
; CURRENT FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: US 60/299,887  
; PRIOR FILING DATE: 2001-06-21  
; PRIOR APPLICATION NUMBER: US 60/301,572  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: US 60/306,501  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: US 60/325,002  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US 60/362,585  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 506  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 425  
; LENGTH: 324  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-177-293-425

Query Match 74.4%; Score 1198; DB 14; Length 324;  
Best Local Similarity 77.3%; Pred. No. 4.4e-108;  
Matches 221; Conservative 27; Mismatches 30; Indels 8; Gaps 2;  
Qy 21 TATLPTGTSKPPSORVPALTGTASNNDLASLSECPVCDFYVLPPILOQCS 73  
Db 37 TISAAGPSSAVPAAAVISGPGGGGAGPVSPOHHELTSLFECPVCDFYVLPPILOQCS 96  
Qy 74 GHLVCSNCRPKLTCCPTCRGLGSIIRNLAMEKVANSVLPFCYASSGCEITLPHTEKAD 132  
Db 97 GHLVCSNCRPKLTCCPTCRGLGSIIRNLAMEKVANSVLPFCYASSGCEITLPHTEKAD 156  
Qy 133 HEELCEPRPSCPCPGASCKWQSLDAMVPHLMHQHKSITTLQGEDIVFLATDINLPGAV 192  
Db 157 HEDICEYRPSCPCPGASCKWQSLDAMVPHLMHQHKSITTLQGEDIVFLATDINLPGAV 216  
Qy 193 DWVMQSCFGFHFMLVLEKQKDYDGHQOFAIVOLIGTRKQAEAFAYRLELNGHRRRLTW 252  
Db 217 DWVMQSCFGFHFMLVLEKQKDYDGHQOFAIVOLIGTRKQAEAFAYRLELNGHRRRLTW 276  
Qy 253 EATPRSHIEGIATAIMNSDCLVFDTSIAQLFAENGNGINVTISM 298  
Db 277 EATPRSHIEGIATAIMNSDCLVFDTSIAQLFAENGNGINVTISM 322

## RESULT 4

US-10-755-889-124  
; Sequence 124, Application US/10755889  
; Publication No. US20040171823A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB  
; FILE REFERENCE: D0284 NP  
; CURRENT APPLICATION NUMBER: US/10/755,889  
; CURRENT FILING DATE: 2004-01-13  
; PRIOR APPLICATION NUMBER: U.S. 60/440,068  
; PRIOR FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: U.S. 60/469,757  
; PRIOR FILING DATE: 2003-05-12  
; NUMBER OF SEQ ID NOS: 823  
; SOFTWARE: Patent In version 3.2  
; SEQ ID NO 124  
; LENGTH: 324





Query Match 25.8%; Score 416; DB 15; Length 324;  
Best Local Similarity 33.8%; Pred. No. 6.8e-32;  
Matches 98; Conservative 44; Mismatches 102; Indels 46; Gaps 10;

QY 35 SORVPALGTATSN-----NDLA-----SLFPCPCVDFYVLPPILOQSGHLVCSNCRPK 84  
DB 31 SSTRPRNGTSTNSVCPNGIAPATSVHELLECPVCTNSMYPPIHQCHNGHTLCTCKTR 90  
QY 85 L-TCCPTCRGRLGSLRNLAMEKVANSVLPFCYASSGCEITLPHTEKADHEELCEFRPYS 143  
DB 91 VHNRCPTCRQELGDIRCALEKVAESLELPCKYISLGCPEIFPYYSKLKHETVCNFRPYS 150  
QY 144 CPCPGASCKQGSLDVAMPHLMHQHKSITTLQGEDIVFLATDNL-----PGAVD-- 193  
DB 151 CPYAGSECVVGDIPFLVAHLRDDHK-----VDMHTGCTFNHRYVKSNPREVENA 200  
QY 194 -WVM-MQSCFGFHFMLVLEKQKVDGHQ-----OFFAIVOLIGTRKOAENFAYRLELNGH 246  
DB 201 TWMLTVFHCFGQYFCL-----HFEAFQLCWAPVYMAFLRPMGDENEARNSYSLEVGAN 254  
QY 247 RRLTWEATPRSIHEGIATATMNSDCLVFDTSIAQLFA--ENGNLGINVT 294  
DB 255 GRKLWEGTFRSDRSHRKVRDSDHGLIIQRNMLFFSGGDRKELKLRVT 304

RESULT 11  
US-10-437-963-111838  
; Sequence 111838, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 111838  
; LENGTH: 693  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_1577C.1.pep  
US-10-437-963-111838

Query Match 25.8%; Score 415; DB 16; Length 693;  
Best Local Similarity 34.7%; Pred. No. 2.3e-31;  
Matches 90; Conservative 50; Mismatches 99; Indels 20; Gaps 8;

QY 48 NNDLASLFEPCVDFYVLPPILOQSGHLVCSNCRPKL-TCCPTCRGRLGSLRNLAMEKV 106  
DB 436 SGNVRELLECPVCLNAMYPIHQCSNGHTLCSGCKPRVHNRCTCRHELGNIRCLALEKV 495  
QY 107 ANSVLPFCYASSGCEITLPHTEKADHEELCEFRPYS CPCPGASCKQGS LDVAMPHLMH 166  
DB 496 AASLELPCKYQNGCLGIYFPYCKLKHESQCYRPTCTPYAGSECTVAGDIQYLVSHLD 555  
QY 167 QHKSITTLQGEDIVFLATDNLPGAVD---WVM-MQSCFGFHFMLVLEKQKVDGHQ--- 219  
DB 556 DHK-VDMHNGSTFNHRYVKS-N-PHEVENATWMLTVFSCFGYFCL-----HFEAFQLCGM 607  
QY 220 --OFFAIVOLIGTRKOAENFAYRLELNGHRRRLTWEATPRSIHEGIATATMNSDCLVFD 277  
DB 608 APVTIAFLRPMGDAAENKNSYSLEVGSGGRKMTQGVPRSIKDRSHKVRDSDYDGLIIQR 667

QY 278 SIAQLFA--ENGNLGINVT 294  
DB 668 NMALFFSGGDKKELKLRVT 686

RESULT 12  
US-10-424-599-161695  
; Sequence 161695, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 161695  
; LENGTH: 311  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_117028C.1.pep  
US-10-424-599-161695

Query Match 25.7%; Score 413.5; DB 15; Length 311;  
Best Local Similarity 32.6%; Pred. No. 1.1e-31;  
Matches 89; Conservative 51; Mismatches 90; Indels 43; Gaps 7;

QY 44 TTASNDLASLFEPCVDFYVLPPILOQSGHLVCSNCRPKL-TCCPTCRGRLGSLRNL 102  
DB 39 TTTSVHD---LLECPVCTNSMYPPIHQCHNGHTLCTCKTRVHNRCTCRQELGDIRCLA 95  
QY 103 MEKVANSVLPFCYASSGCEITLPHTEKADHEELCEFRPYS CPCPGASCKQGS LDVAMP 162  
DB 96 LEKIAESLELPCKYISLGCPEIFPYYSKLKHEAICNFRPYNCYPYAGSDCVSGDIPCLVA 155  
QY 163 HLM-----HGHKSITTLQGEDIVFLATDNLPGAVDWMVMQSCFGFHFMLV 208  
DB 156 HLRRDHRVDMHSGCTFNHRYVKS-NPMEVENATWMLTVFH-----CFGQYFCL- 202  
QY 209 LEKQEKYDGHQ-----OFFAIVOLIGTRKOAENFAYRLELNGHRRRLTWEATPRSIHEGI 263  
DB 203 -----HFEAFQLCWAPVYMAFLRPMGDENEARNSYSLEVGSGGRKLTFFSGSPRSIRDSH 257  
QY 264 ATAIMNSDCLVFDTSIAQLFA--ENGNLGINVT 294  
DB 258 KKVDRSDHGLIIYRNMLFFSGGDRKELKLRVT 290

RESULT 13  
US-10-324-120-1  
; Sequence 1, Application US/10324120  
; Publication No. US20040123349A1  
; GENERAL INFORMATION:  
; APPLICANT: TENASEK LIFE SCIENCES LABORATORY  
; TITLE OF INVENTION: SINAT5, an Arabidopsis thaliana gene involved in lateral root dev  
; FILE REFERENCE: 2577-157  
; CURRENT APPLICATION NUMBER: US/10/324,120  
; CURRENT FILING DATE: 2002-12-20  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-324-120-1

Query Match 25.5%; Score 411; DB 16; Length 309;  
Best Local Similarity 33.3%; Pred. No. 2e-31;

[illegible]

```

RESULT 14
US-10-437-963-158317
  / Sequence 158317, Application US/10437963
  / Publication No. US20040123343A1
  / GENERAL INFORMATION:
  / APPLICANT: La Rosa, Thomas J.
  / APPLICANT: Kovalic, David K.
  / APPLICANT: Zhou, Yihua
  / APPLICANT: Cao, Yongwei
  / APPLICANT: Wu, Wei
  / APPLICANT: Boukharov, Andrey A.
  / APPLICANT: Barbazuk, Brad
  / APPLICANT: Li, Ping
  / TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
  / TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
  / FILE REFERENCE: 38-21(53221)B
  / CURRENT APPLICATION NUMBER: US/10/437,963
  / CURRENT FILING DATE: 2003-05-14
  / NUMBER OF SEQ ID NOS: 204966
  / SEQ ID NO 158317
  / LENGTH: 498
  / TYPE: PRT
  / ORGANISM: Oryza sativa
  / FEATURE:
  / NAME/KEY: unsure
  / LOCATION: (1)..(498)
  / OTHER INFORMATION: unsure at all Xaa locations
  / FEATURE:
  / OTHER INFORMATION: Clone ID: PAT_MRT4530_57802C.1.pcp
  / US-10-437-963-158317

```

[illegible]

```

QY      272 C-LVDETSIAQLFA--ENGNLGINVT 294
      ||:  :|  ||:  :|  :|  :|
Db      453 XGLIIQRNMAFFSGGDRKELKRIT 478

RESULT 15
US-10-324-120-3
; Sequence 3, Application US/10324120
; Publication No. US20040123349A1
; GENERAL INFORMATION:
; APPLICANT: TEWASEK LIFE SCIENCES LABORATORY
; TITLE OF INVENTION: SINAR75, an Arabidopsis thaliana gene involved in lateral root deve
; FILE REFERENCE: 2577-157
; CURRENT APPLICATION NUMBER: US/10/324;120
; CURRENT FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-10-324-120-3

```

Query Match	25.1%	Score 405	DB 16	Length 309
Best local Similarity	33.0%	Pred. No. 7.6e-31		
Matches 93	Conservative 42	Mismatches 93	Indels 54	Gaps 9
Qy	54	LFECPVDFVVLPPILQCSGHLVCSNCRPKL-TCCPTCRGPLSGIRINLAMEKIVANSVLVF	112	
Db	43	LLECPVCTNSMYPPIHQHCHNGYTLCTCKSKRVNKRCTCRQELGDIRCLALEKVAESLEL	102	
Qy	113	PKCYASSGCEITLPHTEKADHEELCEFRPYSCPCFGASCKWQGSIDA WPHLMHOKHSIT	172	
Db	103	PKCYNIIGCGIFIPYYSKLKHESQCNFRPYSCPVAGSECAAVGDTITVAHLRDDHK---	159	
Qy	173	TLOGEDIVFLATDINL-----PGAVD---WVM-MQSCFGFHPMLVLEKQEKYDGHQ-	219	
Db	160	-----VDMHTGCTFNHRYVKSNPVENATWMLTVPCFGQYFCL-----HFEAFQL	206	
Qy	220	-----QFAIVOLIGTRQAEAFAYLELNGHRRRLTWEATPRSIEGIATATMNSDCLVF	275	
Db	207	GMAPVYAFLEFMGDEDDARNYTVSYLEVGSGGRKQTWEGTFPSVDRSHRKVRDSHDGLII	266	
Qy	276	DTSLAQLFIA-----ENGLNGINVTISM C	298	
Db	267	QYNWALFEISGCDKKKELKLVRTGIWKEOONPDSGCIT-SMC	307	

Search completed: April 25, 2005, 06:51:41  
Job time : 404 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 25, 2005, 06:39:49 ; Search time 4911 Seconds  
(without alignments)  
2940.268 Million cell updates/sec

Title: US-10-679-246-2

Perfect score: 1611  
Sequence: 1 MWIIIFLPYVFIEMSRQ.....IAQLFAENGLNINVTISM 298

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO spoal/US10679246/runat\_22042005\_185510\_7685/app\_query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10679246 @CNC 1.1 3731 @runat\_22042005\_185510\_7685 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	1274	6	AR411820 Sequence
2	1611	100.0	1274	6	AX058082 Sequence
3	1611	100.0	2128	9	AK056051 Homo sapi
4	1611	100.0	2440	6	AR380043 Sequence

5	1611	100.0	2324	6	CQ491092	Sequence	
6	1611	100.0	2324	6	CQ493590	Sequence	
7	1611	100.0	2324	6	CQ496955	Sequence	
8	1611	100.0	31705	9	HS4400626	Hom sapi	
9	1611	100.0	173304	9	AC023818	Hom sapi	
10	1603	99.5	2829	6	AX833145	Sequence	
11	1603	99.5	2829	9	AX094663	Hom sapi	
12	1596	99.1	2972	9	HSM807215	Hom sapi	
13	1565	97.1	4090	6	AX780380	Sequence	
C	14	1553.5	96.4	220157	2	AC105462	Rattus no
15	1542	95.7	175636	10	AC142211	Mus muscu	
16	1541.5	95.7	232382	2	AC098991	Rattus no	
17	1540	95.6	1340	9	BC035562	Hom sapi	
18	1540	95.6	2048	9	HS076247	Human hsiAH	
19	1537	95.4	1874	6	CQ719559	Sequence	
20	1537	95.4	2454	9	BC042550	Hom sapi	
21	1536	95.3	1645	10	BC046317	Mus muscu	
22	1536	95.3	1968	6	AX305603	Sequence	
23	1536	95.3	1968	10	MMSIAH1A		
24	1534	95.2	1246	5	BC072747	Xenopus l	
25	1532	95.1	1886	9	HS063295	Hom sapien	
26	1528	94.8	1465	10	AF389476	Rattus no	
27	1527	94.8	2034	9	BC018193	Hom sapi	
C	28	1514.5	94.0	120810	5	BX649644	Zebrafish
29	1514.5	94.0	241301	5	BX470163	Zebrafish	
30	1513	93.9	2419	5	BC045870	Danio rer	
31	1504	93.4	1713	10	MMSIAH1B		
C	32	1504	93.4	112893	10	AL732294	Mouse DNA
C	33	1504	93.4	234105	10	AC091606	Mus Muscu
34	1500	93.1	1720	10	BC052887	Mus muscu	
35	1489	92.4	1884	6	AG3558	Sequence 11	
36	1489	92.4	1884	6	AR271267	Sequence	
37	1486	92.2	1457	10	AB067814	Rattus no	
38	1405	87.2	194622	10	AC133654	Mus muscu	
39	1270	78.8	711	5	CR386772	Gallus ga	
C	40	1262	78.3	168091	10	AC139553	Mus muscu
41	1262	78.3	182733	10	AC123941	Mus muscu	
42	1214	75.4	1441	3	AK112696	Ciona int	
43	1208.5	75.0	2511	10	BC058400	Mus muscu	
44	1206	74.9	2347	9	BC013082	Hom sapi	
45	1205.5	74.8	2048	10	AB067815	Rattus no	

#### ALIGNMENTS

RESULT 1	AR411820	Sequence 1 from patent US 6638734.	1274 bp	DNA	linear	PAT 18-DSC-2003
LOCUS	AR411820	Sequence 1 from patent US 6638734.	1274 bp	DNA	linear	PAT 18-DSC-2003
DEFINITION	AR411820	Sequence 1 from patent US 6638734.	1274 bp	DNA	linear	PAT 18-DSC-2003
ACCESSION	AR411820	Sequence 1 from patent US 6638734.	1274 bp	DNA	linear	PAT 18-DSC-2003
VERSION	AR411820.1	GI:40164258	1274 bp	DNA	linear	PAT 18-DSC-2003
KEYWORDS	AR411820.1	GI:40164258	1274 bp	DNA	linear	PAT 18-DSC-2003
SOURCE	Unknown.	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
ORGANISM	Unknown.	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
REFERENCE	1 (bases 1 to 1274)	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
AUTHORS	Reed, J.C. and Matsuzawa, S.-i.	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
TITLE	Nucleic acid encoding proteins involved in protein degradation,	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
JOURNAL	products and methods related thereto	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
FEATURES	Patent: US 6638734-A 1 28-OCT-2003;	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
source	Location/Qualifiers	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
	1..1274	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
	/organism="unknown"	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003
	/mol_type="genomic DNA"	Unknown.	1274 bp	DNA	linear	PAT 18-DSC-2003

Alignment Scores:	2.79e-121	Length:	1274
Pred. No.:	1611.00	Matches:	298
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	6		

US-10-679-246-2 (1-298) x AR411820 (1-1274)

```
QY 1 MetValIleIleIlePheLeuLeuProProTyValPheIleSerGluMetSerArgGln 20
DB 274 ATGGTTATTAATATTTCTCTGCTCTCTATGATATTTATTTTCAAGAAATGAGCGTCTAG 333
QY 21 ThrAlaThrAlaLeuProThrGlyThrSerTyCysProProSerGlnArgValProAla 40
DB 334 ACTGCTACAGCATACCTACCGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTGCC 393
QY 41 LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys 60
DB 394 CTGACTGGCACAACCTGATCCCAACATGACTTGGCGAGTCTTTTGGAGTGCAGTCTGC 453
QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80
DB 454 TTTGACTATGTTTACCGCCCATCTTCAATGTCCAGAGTGGCCATCTTGTGTAGCAAC 513
QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
DB 514 TGTGCGCCCAAGCTCACATGTTGTCCAACTTGC CGGGGCCCTTTGGGATCCCATTCGCAAC 573
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerGly 120
DB 574 TTTGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 633
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140
DB 634 TTTGAAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 693
QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrIleGlnGlySerLeuAspAlaVal 160
DB 694 CCTTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 753
QY 161 MetProHisLeuMetHisGlnHisSerIleThrThrLeuGlnGlyGluAspIleVal 180
DB 754 ATGCCCCATCTGATGATCAGCATAGCATAGTCATTCATCAACCTACAGGAGGAGGATAGTT 813
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetMetGlnSerCys 200
DB 814 TTTCTTGCTACAGCATTAATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAspGlyHisGlnGln 220
DB 874 TTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAATAATCCATGCTACACAGCAG 933
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240
DB 934 TTTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAAATTTTGTCTTACCGA 993
QY 241 LeuGluLeuAsnGlyHisArgArgArgLeuThrTyrGluAlaThrProArgSerIleHis 260
DB 994 CTTGAGCTAAATGCTCATAGCGCAGCATTCATTCGCTGGAAGCAGCTCTCTGATCTATTTCAT 1053
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280
DB 1054 GAAGGAATTTGCAACAGCCATTAATGATAGGAGCTGCTAGTCTTTGACACACAGCATTTGCA 1113
QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298
DB 1114 CAGCTTTTTCAGAAAATGGCAATTAATAGGCATCAATTAATCAATTTTCCATGCTGT 1167
```

RESULT 2  
AX058082 1274 bp DNA linear PAT 17-JAN-2001  
LOCUS  
DEFINITION Sequence 1 from Patent WO0077207.  
ACCESSION AX058082  
VERSION AX058082.1 GI:12310662  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1  
AUTHORS Reed, J.C. and Matsuzawa, S.I.  
TITLE Nucleic acid encoding proteins involved in protein degradation,  
products and methods related thereto  
JOURNAL Patent: WO 0077207-A 1 21-DEC-2000;  
The Burnham Institute (US)  
FEATURES  
source Location/Qualifiers  
1. .1274  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
274. .1170  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC22474.1"  
/db\_xref="GI:12310663"  
/translation="MWIIIFLLPPYVFI SEMSRQTALPTGTSCKBPSPORVPALGT  
TASNDLASLPECPVDFYLPPLTQCSGHLVCSNCRPKLTCTCTCRGPIGSLRNL  
MEKVANSVLPCKVASSGCEITLPHTEKAHEELCEFRPYSPCPGASCKWQSGSLDA  
MPLHMHOKSITTLQGEDIVFLATDINLPAGVDMWQSCFGPHFMLVLEKQEKYDGH  
QQFAIVQLIGTRKQANFAVRLNGLHRRRLTWEATPERSIHEGIATAMNSDCLFED  
TSLAQLFAENGLNINVTISM"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,79e-121 Length: 1274  
Score: 1611.00 Matches: 298  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 6  
US-10-679-246-2 (1-298) x AX058082 (1-1274)  
QY 1 MetValIleIleIlePheLeuLeuProProTyValPheIleSerGluMetSerArgGln 20  
DB 274 ATGGTTATTAATATTTCTCTGCTCTCTATGATATTTATTTTCAAGAAATGAGCGTCTAG 333  
QY 21 ThrAlaThrAlaLeuProThrGlyThrSerTyCysProProSerGlnArgValProAla 40  
DB 334 ACTGCTACAGCATACCTACCGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTGCC 393  
QY 41 LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys 60  
DB 394 CTGACTGGCACAACCTGCAATCCAACTGACTTGGCGAGTCTTTTGGAGTGCAGTCTGC 453  
QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80  
DB 454 TTTGACTATGTTTACCGCCCATCTTCAATGTCCAGAGTGGCCATCTTGTGTAGCAAC 513  
QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
DB 514 TGTGCGCCCAAGCTCACATGTTGTCCAACTTGC CGGGGCCCTTTGGGATCCCATTCGCAAC 573  
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerGly 120  
DB 574 TTTGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 633  
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140  
DB 634 TTTGAAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 693  
QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrIleGlnGlySerLeuAspAlaVal 160  
DB 694 CCTTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 753  
QY 161 MetProHisLeuMetHisGlnHisSerIleThrThrLeuGlnGlyGluAspIleVal 180  
DB 754 ATGCCCCATCTGATGATCAGCATAGCATAGTCATTCATCAACCTACAGGAGGAGGATAGTT 813  
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetMetGlnSerCys 200



Db 814 TTTCCTGCTACAGACATTAATCTCTCCGTGGTCTGTTCACTGGGTGATGATGCAGTCTCTGT 873  
 Qy 201 PheGlyPheHisPheMetLeuValLeuGlyLeuGlySerGlnGlyTyrAspGlyHisGlnGln 220  
 Db 874 TTGGCTTTTTCACCTTCATGTTAGTCTTAGAGAAAACAGGAAAAATACGATGCTCACCAGCAG 933  
 Qy 221 PhePheAlaIleValGlnLeuIleGlyThrArglySerGlnAlaGluAenPheAlaTyrArg 240  
 Db 934 TTCTTCGCATTCGTACAGCTGTATAGGAACACGCAAGCAAGCTGAAATTTTGTCTACCGA 993  
 Qy 241 LeuGluLeuAenGlyHisArgArgLeuThrTyrGluAlaThrProArgSerIleHis 260  
 Db 994 CTTCAGCTAAATGGTCTATAGCGACGATTGACTTGGGAGCGACTCTCGATCTATTCAT 1053  
 Qy 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280  
 Db 1054 GAAGGAATTCGAACAGCGCAATTATGAATAGCGACTGTCTAGTCTTTGACACCAGCATTCGA 1113  
 Qy 281 GlnLeuPheAlaGluAenGlyAenLeuGlyIleAenValThrIleSerMetCys 298  
 Db 1114 CAGCTTTTTCAGAAAAATGGCAATTTAGGCATCAATGTAATCTATTTCCATGTGT 1167

RESULT 3  
 AK056051/c  
 LOCUS AK056051 2128 bp mRNA linear PRI 30-JAN-2004  
 DEFINITION Homo sapiens cDNA FLJ31489 fis, clone NT2NE2003308.  
 ACCESSION AK056051  
 VERSION AK056051.1 GI:16551141  
 KEYWORDS oligo capping; fis (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1  
 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,  
 Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,  
 Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,  
 Ishii, S., Yamamoto, K., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,  
 Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,  
 Shiratori, A., Sudo, H., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H.,  
 Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,  
 Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,  
 Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,  
 Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,  
 Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,  
 Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,  
 Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,  
 Takeuchi, K., Arita, M., Inose, N., Musashino, K., Yuuki, H., Oshima, A.,  
 Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,  
 Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,  
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,  
 Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,  
 Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,  
 Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,  
 Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,  
 Ozaki, K., Hiroa, M., Ohmori, Y., Kawabata, A., Hiki, T., Kobatake, N.,  
 Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,  
 Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K.,  
 Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,  
 Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,  
 Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,  
 Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,  
 Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.  
 Complete sequencing and characterization of 21,243 full-length  
 human cDNAs  
 Nat. Genet. 36 (1), 40-45 (2004).

2  
 Nidomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,  
 Kodaira, H., Furiya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,  
 Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,  
 Yamamoto, T., Irie, R., Otsuki, T., Satoh, H., Wakamatsu, A., Ishii, S.,  
 Sugiyama, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,

TITLE  
 JOURNAL  
 PUBMED  
 14702039  
 REFERENCE  
 AUTHORS

	Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuko,Y., Nagai,K. and Isogai,T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 2128)
AUTHORS	Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE	Direct Submission
JOURNAL	Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1539-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@ri.co.jp, Tel.81-438-52-3975, Fax:81-438-52-3986)
COMMENT	(NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES	Location/Qualifiers 1..2128 <code>/organism="Homo sapiens"</code> <code>/mol_type="mRNA"</code> <code>/db_xref="taxon:9606"</code> <code>/clone="NT2NE2003308"</code> <code>/cell_line="NT2"</code> <code>/cell_type="teratocarcinoma"</code> <code>/clone_lib="NT2NE2"</code> <code>/note="cloning vector: pME18SFL3-mRNA from NT2 neuron after the differentiation of NT2 neuronal precursor cells."</code>
ORIGIN	
Alignment Scores:	
Pred. No.:	5,17e-121 Length: 2128
Score:	1611.00 Matches: 298
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	Gaps: 0
US-10-679-246-2 (1-298) x AK056051 (1-2128)	
Qy	1 MetValIleIlePheLeuLeuProProTyrValPheIleSerGluMetSerArgGln 20
Db	2025 ATGGTTATAATTATTTTTTCCTCGCCTCCCTATGATTATTTTCAGAAATGAGCCGTGAC 1966
Qy	21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSerGlnArgValProAla 40
Db	1965 ACTGCTACAGCATTAACCTACCGGTACTCGAAGTGTCACCATCCAGAGGGTGCCTGCC 1906
Qy	41 LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys 60
Db	1905 CTGACTGGCACAACTGCATCCAAACAATGACTTGGCGAGTCTTTTGGTGTCCAGTCTGC 1846
Qy	61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80
Db	1845 TTTGACTATGTGTTCGCGCCATCTCTCAATGTCAGATGGCCATCTTGTTGTAGAACA 1786
Qy	81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
Db	1785 TGTGCGCCCAAAGCTCACATGTTGTCCAATTGCGGGGCCCTTTGGGATCCATTGCGNAAC 1726
Qy	101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120
Db	1725 TTGGCTATGAGAAAAGTGCGCTAAATCAGTACTTTTCCCCTGTAAAAATGCGTCTTCGGA 1666
Qy	121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140
Db	1665 TGTTAAATAACTCTGCCACACACAGAAAANGACACATGAGAGCTCTGTGAGTTTAGG 1606
Qy	141 ProTyrSerCysProCysProGlyAlaSerCysLysTrpGlnGlySerLeuAspAlaVal 160

Db	1605	CCTTATTTCCTGTCCTCGGCCCTGGTCTCTCTGTAATGGCAAGGCTCTCTCGATGCTGTA	1546
Qy	161	MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal	180
Db	1545	ATGCCCACTGATGCATCAGCATTAAGTCCATTACAACCTACAGGAGAGGATATAGTT	1486
Qy	181	PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSerCys	200
Db	1485	TTTCTTGCTACACACATTATATCTTCTGGTGCTGTGACTGGGTGATGATGCAGTCTCTGT	1426
Qy	201	PheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAspGlyHisGlnGln	220
Db	1425	TTTGGCTTTCACCTTCATCTTAGTCTTAGAGAACACAGGAAAAATACGATGGTCAACAGCAG	1366
Qy	221	PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg	240
Db	1365	TTCTTCGCAATCGTCAGCTGATAGGAACACGCAAGCAAGCTGAAAAATTGTCTTACCGA	1306
Qy	241	LeuGluLeuAsnGlyHisArgArgArgLeuThrTyrGluAlaThrProArgSerIleHis	260
Db	1305	CTTGAGCTAATGGTCATAGCGACGATTCGACTTGGAGCGACTCTCGATCTATTCTAT	1246
Qy	261	GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla	280
Db	1245	GAAGGAATTCGAACAGCCATTATGATAGCGACTGTCTAGTCTTTGACACACAGCATTGCA	1186
Qy	281	GlnLeuPheAlaGluAsnGlyValLeuGlyIleAsnValThrIleSerMetCys	298
Db	1185	CAGCTTTTTCAGAAAAATGGCAATTATAGGCATCAATGAACTATTTCCATGTGT	1132
RESULT 4			
AR380043	AR380043	2440 bp	DNA
LOCUS	Sequence 588 from patent US 6607879.	linear	PAT 18-DEC-2003
DEFINITION	Sequence 588 from patent US 6607879.		
ACCESSION	AR380043		
VERSION	AR380043.1	GI:40087677	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2440)		
AUTHORS	Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.		
TITLE	Compositions for the detection of blood cell and immunological response gene expression		
JOURNAL	Patent: US 6607879-A 588 19-AUG-2003;		
FEATURES	Location/Qualifiers		
source	1..2440		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:			
Pred No.:	6, 1e-121	Length:	2440
Score:	1611.00	Matches:	298
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-679-246-2 (1-298) x AR380043 (1-2440)			
Qy	1	MetValIleIleIlePheLeuLeuProTyrValPheIleSerGluMetSerArgGln	20
Db	507	ATGGTATAAATTATTTTCTCTCGCTCCTTANGATTATTTTTCAGAAATGAGCGGTGAG	566
Qy	21	ThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSerGlnArgValProAla	40
Db	567	ACTGCTACAGCATTACCTACCGGTACTTCGAAGTGTCCACCATCCACAGAGGTGCTGCC	626
Qy	41	LeuThrGlyThrThrAlaSerAsnAsnAspIleAlaSerLeuPheGluCysProValCys	60
Db	627	CTGACTGGCACAACATGCATCCAACTAATGACTTGGCGAGTCTTTTGGTGTCCAGTCTGC	686



Db 1353 TTGGCTTTCACTTCATGTTAGTCTTAGAGAAAACAGGAAAAATACGATGGTCACCAGCAG 1412

Qy 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240

Db 1413 TTCTTCGCAATGTCACAGCTGATAGCAACAGCAAGCTGAAAAATTTTGGCTTACCAG 1472

Qy 241 LeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260

Db 1473 CTTGAGCTAAATGGTCATAGGCGAGATTCGCTTGGGAAAGGAGCTCTTCGATCTATTTCAT 1532

Qy 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280

Db 1533 GAAGGAATTCGCAACAGCCATTATGATAGCGACTGTCTAGCTTTTGACACAGCATTGCA 1592

Qy 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298

Db 1593 CAGCTTTTTCGAGAAAATGGCAATTTAGGCATCAATGTAATCTATTTCCTATGTGT 1646

RESULT 7

LOCUS CQ496955 2924 bp DNA linear PAT 30-JAN-2004

DEFINITION Sequence 28822 from Patent WO0160860.

ACCESSION CQ496955

VERSION CQ496955.1 GI:41462591

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 Schlegel, R., Endege, W.O. and Monahan, J.E.

AUTHORS use

TITLE Genes differentially expressed in human prostate cancer and their

JOURNAL Patent: WO 0160860-A 28822 23-AUG-2001;

Millennium Predictive Medicine, Inc. (US)

FEATURES

source

1. .2924

/organism="Homo sapiens"

/db\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 7 58e-121 Length: 2924

Score: 1611.00 Matches: 298

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-679-246-2 (1-298) x CQ496955 (1-2924)

Qy 1 MetValIleIleIlePheLeuLeuProProTyrValPheIleSerGluMetSerArgGln 20

Db 753 ATGGTTATAATTTTCTCTCTGCTCTTATGTTATTTATTTTTCAGAAATGAGCCGTGAC 812

Qy 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla 40

Db 813 ACTGCTACAGATTACTACCGTACTCGAGTGTCCACCATCCAGAGGGTGCCTGCC 872

Qy 41 LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys 60

Db 873 CTGACTGGCACAACGTGCATCAACAAATGACTTGGCGAGTCTTTTGTAGTGTCCAGCTGC 932

Qy 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80

Db 933 TTTGACTATGTGTTACCGCCCAATTCCTCAATGTACAGAGTGCCCAATCTTTGTTGTAGCAAC 992

Qy 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100

Db 993 TGTGCGCCCAAGCTCACATGTTGTCCAACTTCGCGGGGCCCTTTGGGATCCATTCGCAAC 1052

Qy 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120

Db 1053 TTGGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGGA 1112

Qy 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140

Db 1113 TGTGAATAAATCTTCGCCACACACAGAAAAAGACACCATCAAGAGCTCTGTGAGTTAGG 1172

Qy 141 ProTyrSerCysProCysProGlyAlaSerCysLysTrpGlnGlySerLeuAspAlaVal 160

Db 1173 CCTATTCTCTGTCGTCGCTCTGCTTCTCTGTAATGGCAAGGCTCTCTGGATGCTGTA 1232

Qy 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180

Db 1233 ATGCCCATCTGATGTCATCAGCATAAGTCCATTACAACCCCTACAGGGAGAGGATATAGTT 1292

Qy 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSerCys 200

Db 1293 TTTCTTGTCTCAGACATTAATCTTCTCTGCTGCTCTGTAATGGGTCATGATGCACTCTGT 1352

Qy 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAspGlyHisGlnGln 220

Db 1353 TTTGGCTTTCACTTCATGTTAGTCTTAGAGAAAACAGGAAAAATACGATGGTCACCAGCAG 1412

Qy 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240

Db 1413 TTTCTTGCATATCGTACAGCTGATAGCAACAGCAAGCAAGCTGAAAAATTTTGTCTTACCAG 1472

Qy 241 LeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260

Db 1473 CTTGAGCTTAAATGGTTCATAGCGAGCTGATGCTTGGGAAGCGACTCTCTCGATCTATTTCAT 1532

Qy 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280

Db 1533 GAAGGAATTCGCAACAGCCATTATGATAGCGACTGTCTAGCTTTTGACACAGCATTGCA 1592

Qy 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298

Db 1593 CAGCTTTTTCGAGAAAATGGCAATTTAGGCATCAATGTAATCTATTTCCTATGTGT 1646

RESULT 8

LOCUS HSA400626 31705 bp DNA linear PRI 02-APR-2001

DEFINITION Homo sapiens SIAH1 gene, exons 1-2.

ACCESSION AJ400626

VERSION AJ400626.1 GI:13539602

KEYWORDS siah1 gene.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Medhioub, M., Muchardt, C., Tubacher, E., Giudicelli, C., Hors-Cayla, M.C. and Thomas, G.

TITLE Down regulation of the TATA-less and GC-rich SIAH1 promoter by TP53

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 31705)

AUTHORS Medhioub, M.

TITLE Direct Submission

JOURNAL Submitted (11-APR-2000) Medhioub M., Fondation Jean DAUSSET. CEPH, 27, rue Juliette Dodu, Paris 75010, FRANCE

FEATURES

source

1. .31705

/organism="Homo sapiens"

/db\_type="genomic DNA"

/db\_xref="taxon:9606"

/map="16q12-q13"

5126. .30193

/gene="SIAH1"

5126. .5550

/number=1

/evidence=experimental

5126. .5550

```

/feature="SIAH1"
/evidence=experimental
5551..28322
/feature="SIAH1"
/evidence=experimental
28323..30193
/feature="SIAH1"
/number=2
/evidence=experimental
28325..29173
/feature="SIAH1"
/codon start=1
/product="SIAH1 protein"
/protein_id="CAC35342.1"
/db_xref="GI:13539603"
/db_xref="GOA:O43269"
/db_xref="UniProt/TREMBL:O43269"
/translation="MSRQATALPTGTSKPSQRPALPTGTASNDLASFPECPVC
FYVLPILQCSGHLVCSNCRPKLTCTCTCRGPIGSIRNLAMKVANSVLFPCKYAS
SGCEITLPTEKADHEELCEFPYPCPGKSGKQWGLSDAVPHMLHQHKSITTLQ
EDIVFLATDINLPFADVWMQSCFGFHEMLVLEKQYDGHQFPFAIVQLIGTRKQA
ENFAYRLNGLHGRRLTWEATPSRIHEGIATAIMNSDCLVFDTSTIAQLFAENGNLGIN
VTISM"
29173..30193
/feature="SIAH1"
/evidence=experimental
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 1,34e-119 Length: 31705
Score: 1611.00 Matches: 298
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-679-246-2 (1-298) x HSA400626 (1-31705)
QY 1 MetValIleIleIlePheLeuLeuProProTyValPheIleSerGluMetSerArgGln 20
DB 28277 ATGGTTATAATATTTTCTCTGCTCTCTTATGATATTATTTCAGAAATGAGCGCTCAG 28336
QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSerGlnArgValProAla 40
DB 28337 ACTGCTACAGCATACCTACCGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTGCC 28396
QY 41 LeuThrGlyThrThrAlaSerIleAsnAspLeuAlaSerLeuPheGluCysProValCys 60
DB 28397 CTGACTGGCACCACTGCATCCCAATGACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGC 28456
QY 61 PheAspTyThrValLeuProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80
DB 28457 TTTGACTATGTGTACCGCCCATCTTCAATGTCAGAGTGGCCATCTTGTGTAGCAAC 28516
QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
DB 28517 TGTGCGCCCAAGCTACATGTTGTCTCAACTGTCGCGGGCCCTTTGGGATCCATTCGCAAC 28576
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyAlaSerSerGly 120
DB 28577 TTGGCTATGAGAAAGTGCTTAATCAGTACTTTTCCCTGTAATATGCTTCTTCGGA 28636
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140
DB 28637 TGTGAATAACTCTGCCACACAGAAAGCAGACCATGAGAGCTCTGTGAGTTAGG 28696
QY 141 ProTySerCysProCysProGlyAlaSerCysLysTyTrpGlnGlySerLeuAspAlaVal 160
DB 28697 CCTTATCTCTGTCGTCCTGTGTCTCTCTGTAATGCAAGGCTCTCTGATGCTGTA 28756
QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180
DB 28757 ATGCCCCCATCTGATGCATAGCATAGTCCATTACAACCCCTACAGGAGATATAGTT 28816

```

```

QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSerCys 200
DB 28817 TTTCTTGTCTACAGACATTAACTCTCTGCTGCTGTTGACTGGGTGATGATGCAGTCTGT 28876
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyAspGlyHisGlnGln 220
DB 28877 TTTGGCTTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 28936
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyArg 240
DB 28937 TTTCTTGGCAATCGTACAGCTGATAGGAACACGCAAGCAGCTGAAATTTTGTCTTACCGA 28996
QY 241 LeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260
DB 28997 CTTGAGCTAAATGGTCATAGGCGCAGCATTTGACTTGGGAAGCAGCTCCTCGATCTATTCAT 29056
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280
DB 29057 GAAGGAATTCGAACACGCCATTATGAATAGCAGCTGTCTAGTCTTTGACACCGCATTCGA 29116
QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298
DB 29117 CAGCTTTTTCAGAAATGGCAATTTAGGCATCAATGTAATCTATTTCCATGTGT 29170
RESULT 9
AC023818 173304 bp DNA linear PRI 03-JAN-2004
LOCUS Homo sapiens chromosome 16 clone CTD-2600H12, complete sequence.
AC023818
VERSION AC023818.5 GI:40556309
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 173304)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 173304)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 173304)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-MAY-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 173304)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2004) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jan 3, 2004 this sequence version replaced gi:20429291.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
Location/Qualifiers
1..173304
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTD-2600H12"
FEATURES
source
ORIGIN

```

LOCUS AX833145 2829 bp DNA linear PAT 15-DEC-2003  
DEFINITION Sequence 269 from Patent EPI1347046.  
ACCESSION AX833145  
VERSION AX833145.1 GI:39919280  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuho, Y.  
TITLE Full-length cDNA sequences  
JOURNAL Patent: EP 1347046-A 269 24-SEP-2003;  
RESEARCH Association for Biotechnology (JP)  
FEATURES  
source location/Qualifiers  
1. 2829  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.24e-120 Length: 2829  
Score: 1603.00 Matches: 297  
Percent Similarity: 99.66% Conservativity: 0  
Best Local Similarity: 99.66% Mismatches: 1  
Query Match: 99.50% Indels: 0  
DB: Gaps: 0  
US-10-679-246-2 (1-298) x AX833145 (1-2829)  
QY 1 MetValIleIlePheLeuLeuProTyrValPheIleSerGluMetSerArgGln 20  
DB 1407 ATGGTTAATTAATTTTCTCTGCTCTTATGTTATTTATTCAGAAATGAGCGCTCAG 1466  
QY 21 ThrAlaThrAlaLeuProThrGlyThrSerIysCysProSerGlnArgValProAla 40  
DB 1467 ACTGCTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTGCC 1526  
QY 41 LeuThrGlyThrThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCys 60  
DB 1527 CTGACTGGCACAACCTGCAATCAACAAATGACTTGGCGAGTCTTTTGGAGTCCAGTCTGC 1586  
QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80  
DB 1587 TTTGACTATGTGTACCGCCCATTTCTCAATGTCAGAGTGGCCATCTGTTTGTAGCAAC 1646  
QY 81 CysArgProIysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
DB 1647 TGTGCGCCCAAGCTCACATGTTGTCCAACTTGGCGGGCCCTTTGGGATCCCAATTCGCAAC 1706  
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
DB 1707 TTGGCTATGGAGAAAGTGGCTTAATTCAGTACTTTTCCCTGTAAATATCGCTTCTGGA 1766  
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluLeuLeuCysGluPheArg 140  
DB 1767 TGTGAAATAACTCTGCCACACAGAAAAAGCAGACCATGAAGAGCTCTGTGTAGTTAG 1826  
QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal 160  
DB 1827 CCTATTCTCTCGCGCCCTGGTCTCTCTGTAAATGCAAGGCTCTCTGGATGCTGTA 1886  
QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180  
DB 1887 ATGCCCCCATCTGATGATCAGCATAGTCCATTAACCAACCTACAGGAGAGGATATAGTT 1946  
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetMetGlnSerCys 200  
DB 1947 TTTCTTGTACAGACATTAACTTCTCTGTGTGTGTGTTGACTGGGTGATGATGTCAGTCTGT 2006  
RESULT 10  
AX833145







gene	CDS
g00001	g00001
g00002	g00002
g00003	g00003
g00004	g00004
g00005	g00005
g00006	g00006
g00007	g00007
g00008	g00008
g00009	g00009
g00010	g00010
g00011	g00011
g00012	g00012
g00013	g00013
g00014	g00014
g00015	g00015
g00016	g00016
g00017	g00017
g00018	g00018
g00019	g00019
g00020	g00020
g00021	g00021
g00022	g00022
g00023	g00023
g00024	g00024
g00025	g00025
g00026	g00026
g00027	g00027
g00028	g00028
g00029	g00029
g00030	g00030
g00031	g00031
g00032	g00032
g00033	g00033
g00034	g00034
g00035	g00035
g00036	g00036
g00037	g00037
g00038	g00038
g00039	g00039
g00040	g00040
g00041	g00041
g00042	g00042
g00043	g00043
g00044	g00044
g00045	g00045
g00046	g00046
g00047	g00047
g00048	g00048
g00049	g00049
g00050	g00050
g00051	g00051
g00052	g00052
g00053	g00053
g00054	g00054
g00055	g00055
g00056	g00056
g00057	g00057
g00058	g00058
g00059	g00059
g00060	g00060
g00061	g00061
g00062	g00062
g00063	g00063
g00064	g00064
g00065	g00065
g00066	g00066
g00067	g00067
g00068	g00068
g00069	g00069
g00070	g00070
g00071	g00071
g00072	g00072
g00073	g00073
g00074	g00074
g00075	g00075
g00076	g00076
g00077	g00077
g00078	g00078
g00079	g00079
g00080	g00080
g00081	g00081
g00082	g00082
g00083	g00083
g00084	g00084
g00085	g00085
g00086	g00086
g00087	g00087
g00088	g00088
g00089	g00089
g00090	g00090
g00091	g00091
g00092	g00092
g00093	g00093
g00094	g00094
g00095	g00095
g00096	g00096
g00097	g00097
g00098	g00098
g00099	g00099
g00100	g00100

QY 241 LeuGluLeuAsnGlyHisArgArgLeuThrTTPGluAlaThrProArgSerIleHis 260  
 Db 1734 CTTGAGCTAAATGGTTCATAGGCGACGATTGACTTTGGGAAGCGGACTCTCGATCTATTTCAT 1793  
 QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280  
 Db 1794 GGAGGAAATGCAACGCCATTATGATAGCGACTGTCTAGTCTTTGACACCGCAGCATGCA 1853  
 QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298  
 Db 1854 CAGCTTTTTCGAGAAAATGCAATTTAGGCATCAATGTAATCAATTTCCATGTGT 1907  
 RESULT 13  
 LOCUS AX780380/c 4090 bp DNA linear PAT 14-JUL-2003  
 DEFINITION Sequence 2537 from Patent WO03039443.  
 ACCESSION AX780380  
 VERSION AX780380.1 GI:32697374  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
 Dugas, M., Ellis, R., Brors, B. and Mergenthaler, S.  
 TITLE Novel genetic markers for leukemias  
 JOURNAL Patent: WO 03039443-A 2537 15-MAY-2003;  
 Deutsches Krebsforschungszentrum (DE);  
 Ludwig-Maximilian-Universitaet Muenchen (DE); Haferlach, Torsten,  
 PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)  
 FEATURES  
 source  
 1. 4090  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6,09e-117 Length: 4090  
 Score: 1565.00 Matches: 295  
 Percent Similarity: 98.33% Conservative: 0  
 Best Local Similarity: 98.33% Mismatches: 3  
 Query Match: 97.14% Indels: 2  
 DB: Gaps: 0  
 US-10-679-246-2 (1-298) x AX780380 (1-4090)  
 QY 1 MetValIleIlePheLeuLeuProProTyrValPheIleSerGluMetSerArgGln 20  
 Db 3338 ATGGTTATTAATTAATTTCTCTGCTCCTCTATGTATTTATTTTCAGAAATGAGCGCTCAG 3279  
 QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSerGlnArgValProAla 40  
 Db 3278 ACTGCTACAGCATTAATACCGTACTCTGAAGTGTCACCATCCAGAGGWTGCTGCC 3219  
 QY 41 LeuThrGlyThrThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCys 60  
 Db 3218 CTGACTGGCACAACATCATCAACATGACTTGGCGAGTCTTTTGAGTGTCAGCTGCG 3159  
 QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80  
 Db 3158 TTTGACTATGTGTATTCGCGCCCATTTCTCAATGTTCAGATGGCCATCTTGTGTAGCAAC 3099  
 QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
 Db 3098 TGTGCGCCCAAGCTCATGTGTTCACATGTGCGGGGCCCTTTGGGATCCATTCGCAAC 3039  
 QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
 Db 3038 TTGGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATATCGTCTTCTGA 2979  
 QY 121 CysGluIleThrLeuProHisThrGluLys-AlaAspHisGlu-GluLeuCysGluPheA 140

Db 2978 TGTGAATAAATCTCTGNACACACAGAAAAAANGAGACCACCATGAANANGCTCTGTGAGTTTA 2919  
 QY 140 rGProTyrSerCysProCysProGlyValaSerCysLysTyrGlnGlySerLeuAspAlaV 160  
 Db 2918 GGCTTATTCTGTGTCGTCGCCCTGGTGCTCTCTGTAATGSCAAGGCTCTCTGATGCTG 2859  
 QY 160 alMerProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleV 180  
 Db 2858 TAATGCCCATCTGATGCATCAGCATAGTCCATTACAACCCCTACAGGGAGAGATATAG 2799  
 QY 180 alPheLeuAlaThrAspIleAsnLeuProGlyValaValaAspTyrValMetMetGlnSerC 200  
 Db 2798 TTTTCTTCTCTACAGACATTAATCTTCTGGTGCTGTGACTGGGTGATGATCCAGTCCCT 2739  
 QY 200 yPheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAspGlyHisGlnG 220  
 Db 2738 GTTTTGGCTTTTCATCTTCTAGTCTGTAGAGAAACAGGAAATAACGATGGTCCACCAGC 2679  
 QY 220 InPhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrA 240  
 Db 2678 AGTTCTTCGCAATCGTACAGCTGTAGAGAACACGCAAGCAAGCTGAAAAATTTTGTCTACC 2619  
 QY 240 rGLeuGluLeuAsnGlyHisArgArgLeuThrTTPGluAlaThrProArgSerIleH 260  
 Db 2618 GACTTGAGCTAAATGGTTCATAGCGCAGGATTCGTTGGGAAGCGACTCTCTCGATCTATT 2559  
 QY 260 isGluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleA 280  
 Db 2558 ATGAAGGAATTCACACAGCCATTATGATAGCGACTGTCTAGTCTTTGACACAGCATTCG 2499  
 QY 280 laGlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298  
 Db 2498 CACAGCTTTTTCGAGAAAATGCAATTTAGGCATCAATGTAATCTTTCCATGTGT 2443  
 RESULT 14  
 AC105462  
 LOCUS Rattus norvegicus clone CH230-145D2, WORKING DRAFT SEQUENCE.  
 DEFINITION AC105462  
 ACCESSION AC105462  
 VERSION AC105462.7 GI:30521365  
 HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
 KEYWORDS Rattus norvegicus (Norway rat)  
 SOURCE Rattus norvegicus  
 ORGANISM Mammalia; Eutheria; Rodentia; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Metazoa; Chordata; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 220157)  
 AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abranzon, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
 Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., J.,  
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorensuhewa, L., Louiseged, H., Lozado, R. J., Lu, X., Ma, J.,

\* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 220157: contig of 220157 bp in length.

# FEATURES

source  
 1. 220157  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-145D2"  
 1. 997  
 /note="wgs end extension  
 clone end:T7"  
 1295..2037  
 /note="clone boundary  
 clone end:T7"  
 site:EcORI  
 end\_sequence:BH297589"  
 misc\_feature  
 misc\_feature

# ORIGIN

Alignment Scores:  
 Pred. No.: 6.31e-114 Length: 220157  
 Score: 1553.50 Matches: 286  
 Percent Similarity: 97.67% Conservative: 8  
 Best Local Similarity: 95.02% Mismatches: 4  
 Query Match: 96.43% Indels: 3  
 Gaps: 1  
 DB: 2

US-10-679-246-2 (1-298) x AC105462 (1-220157)

QY 1 MetValIleIleIlePheLeuLeuProProTyValPheIle-----SerGluMet 17  
 Db 1181118 GTTATATTGTAACCTTCCCTGCTCTCTTTATTTCTTTTCTTACAGAGATG 118177  
 QY 18 SerArgGlnThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArg 37  
 Db 118178 AGCCGTGACAGCTGTACAGCATTACCCACCGGCACCTCAAGTGTCCACCATCCAGAGG 118237  
 QY 38 ValProAlaLeuThrGlyThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCys 57  
 Db 118238 GTACCTGCTTGCCTTGCCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 118297  
 QY 58 ProValCysPheAspTyValLeuProProIleLeuGlnCysGlnSerGlyHisVal 77  
 Db 118298 CCTGCTGCTTGTGACTATGTTATGTCACCTATTTCTTCAGTGTTCAGAGTGCCATCTGT 118357  
 QY 78 CysSerAsnCysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySer 97  
 Db 118358 TGTAGCACTGTGCGCCCAACTTACATGTTGTCCACCTGCGGGGCGGCGTGGATCC 118417  
 QY 98 IleArgAsnLeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyAla 117  
 Db 118418 ATTGCAACTTGGCTATGAGAAAGTGGCAACTCGTACTCTTCTTGTAAATATGCC 118477  
 QY 118 SerSerGlyCysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCys 137  
 Db 118478 TCTTCTGGATGTGAGATAACTCTGCGGCACACCGAAAGCAGAGCAGGAGGCTCTGT 118537  
 QY 138 GluPheArgProTySerCysProCysProGlyAlaSerCysLysTrpGlnGlySerLeu 157  
 Db 118538 GAGTTGAGGCTTACTTCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 118597  
 QY 158 AspAlaValMetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGlu 177  
 Db 118598 GATGCTGTCAATGCGCCCACTGATGATGATGATGATGATGATGATGATGATGATGATG 118657  
 QY 178 AspIleValPheLeuAlaThrAspIleAsnLeuProGlyValaValAspTrpValMetMet 197  
 Db 118658 GATATAGTTTCTTGTCTACAGACATTAACTTCTTGTGCTGCTGCTGCTGCTGCTGCTG 118717  
 QY 198 GlnSerCysPheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyAspGly 217

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwaokemele, O., Okunolu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfanckoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

# TITLE

# JOURNAL

# REFERENCE

# AUTHORS

# TITLE

# JOURNAL

# REFERENCE

# AUTHORS

# TITLE

# JOURNAL

# COMMENT

Submitted (09-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 Rat Genome Sequencing Consortium.  
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On May 10, 2003 this sequence version replaced gi:25008590.  
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

# ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GHPM

Center clone name: CH230-145D2

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 215295 bases at least Q40

Consensus quality: 216317 bases at least Q30

Consensus quality: 217167 bases at least Q20

Estimated insert size: 223095; sum-of-contigs estimation

Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 1 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

```
Db 118718 CAGTCTGTTTGGCTTTCATTTCATGTTAGTCTTGGAGAAACAGAAATACGATGGC 118777
|||
QY 218 HisGlnGlnPhePheAlaLeValGlnLeuIleGlyThrArglyGlnAlaGluAnPhe 237
|||
Db 118778 CATCAGCAGTCTTTCGCGATTGTACAGCTGATAGGAACAGCAAGCTGAAATTTT 118837
|||
QY 238 AlaTyArgLeuGluLeuAenGlyHisArgArgLeuThrTrpGluAlaThrProArg 257
|||
Db 118838 GCATATCGACTTTCAGCTGAATGTCTATAGCGCGGATTCGCTTGGGAAGCCACCCCTCG 118897
|||
QY 258 SerIleHisGluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThr 277
|||
Db 118898 TCTATTTCATGAGGCAATTCGACAGGCATCATGATAGTACTGCTAGTGTGACACC 118957
|||
QY 278 SerIleAlaGlnLeuPheAlaGluAenGlyAsnLeuGlyIleAsnValThrIleSerMet 297
|||
Db 118958 AGCATTGCACAGCTCTTTCGAGAAATGCGCAATTAGGCATCAATGTAATATTTCATG 119017
|||
QY 298 Cys 298
|||
Db 119018 TGT 119020

RESULT 15
AC142211 175636 bp DNA linear ROD 10-JUL-2004
LOCUS Mus musculus BAC clone RP24-501B23 from chromosome 8, complete
DEFINITION sequence.
ACCESSION AC142211
VERSION AC142211.3 GI:48675474
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 175636)
Cotton,M., Haakenson,W. and Shahid,S.
The sequence of Mus musculus BAC clone RP24-501B23
Unpublished (2001)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 175636)
McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 175636)
Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 175636)
Wilson,R.K.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2004) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jun 12, 2004 this sequence version replaced gi:29244764.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submission@wustl.edu
----- Summary Statistics
Center project name: M_BB0501B23
-----
NOTICE:
This sequence was finished as follows unless otherwise noted:
```

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

#### SOURCE INFORMATION:

The RP24-501 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

#### NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

#### FEATURES

##### Location/Qualifiers

1..175636  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="8"  
/map="8"  
/clone="RP24-501B23"  
/clone\_lib="RP24-501B23"

##### repeat\_region

5..754  
/rpt\_family="L1"

##### repeat\_region

898..1045  
/rpt\_family="B4"

##### repeat\_region

1156..1421  
/rpt\_family="L1"

##### tRNA

2845..2917  
/product="tRNA-Ser"

/note="Likely pseudogene (HMM Sc=34.44 / Sec struct

##### repeat\_region

2848..3027  
/rpt\_family="B2"

##### repeat\_region

3043..3183  
/rpt\_family="L1"

##### repeat\_region

3901..4186  
/rpt\_family="MaLR"

##### repeat\_region

4637..4953  
/rpt\_family="MaLR"

##### repeat\_region

4958..5210  
/rpt\_family="MaLR"

##### repeat\_region

5211..5580  
/rpt\_family="MaLR"

##### repeat\_region

5563..5684  
/rpt\_family="MaLR"

##### repeat\_region

5685..5776  
/rpt\_family="MaLR"

##### repeat\_region

5708..5780  
/rpt\_family="Alu"

##### repeat\_region

6154..6358  
/rpt\_family="B2"

##### repeat\_region

6489..6669  
/rpt\_family="B2"

##### repeat\_region

6671..6864  
/rpt\_family="B2"

##### repeat\_region

6865..6878  
/rpt\_family="B4"

##### repeat\_region

6879..7027  
/rpt\_family="Alu"

##### repeat\_region

7028..7088  
/rpt\_family="B4"

##### repeat\_region

7680..7883  
/rpt\_family="B2"

##### repeat\_region

8256..8757

```

/rpt_family="ERV1"
9203..9293
/rpt_family="MIR"
10285..10327
/rpt_family="MIR"
10331..10827
/rpt_family="ERVK"
11815..12024
/rpt_family="MaLR"
12030..12203
/rpt_family="B2"
12320..12669
/rpt_family="MaLR"
12730..12839
/rpt_family="Alu"
12897..13082
/rpt_family="MaLR"
13526..13593
/rpt_family="MaLR"
13750..13925
/rpt_family="MaLR"
13941..14158
/rpt_family="B2"
14436..14584
/rpt_family="L1"
15670..15743
/rpt_family="ID"
16031..16083
/rpt_family="MIR"
16332..16525
/rpt_family="B4"
16578..16719
/rpt_family="Alu"
16829..17245
/rpt_family="L2"
17253..17350
/rpt_family="B4"
17578..17727
/rpt_family="L2"
17982..18175
/rpt_family="B2"
18489..18769
/rpt_family="ERVK"
18987..19304
/rpt_family="MaLR"
19694..19900
/rpt_family="B2"
20140..20268
/rpt_family="Alu"
20369..20579
/rpt_family="MaLR"
20599..20725
/rpt_family="Alu"
21080..21149
/rpt_family="ID"
21348..21457
/rpt_family="Alu"
22354..22573
/rpt_family="B2"
22986..23157
/rpt_family="MaLR"
23168..23278
/rpt_family="B4"
23456..23520
/rpt_family="MaLR"
23996..24325
/rpt_family="MaLR"
25121..25290
/rpt_family="B4"
25402..25492
/rpt_family="Alu"
25696..25886
/rpt_family="L1"

```

```

repeat_region 26091..26240
/rpt_family="L1"
repeat_region 26241..26382
/rpt_family="Alu"
repeat_region 26383..26516
/rpt_family="L1"

```

# Alignment Scores:

Pred. No.: 4.11e-113 Length: 175636  
 Score: 1542.00 Matches: 286  
 Percent Similarity: 97.96% Conservative: 2  
 Best Local Similarity: 97.28% Mismatches: 4  
 Query Match: 95.72% Indels: 4  
 DB: 10 Gaps: 1

US-10-679-246-2 (1-298) x AC142211 (1-175636)

QY	5	IlePheLeuLeuProProTyrValPheIleSerGluMetSerArgGlnThrAla	24
DB	96935	ATTTTCTTTGTTT-----TTCTTTACAGAGATGAGCCGACACTGTACAGCA	96982
QY	25	LeuProThrGlyThrSerLysCysProProSerGlnArgValProAlaLeuThrGlyThr	44
DB	96983	TTACCCACTGGCACCCTCAAGTGTCCACCATCCACAGAGGTACCTGCTTGACCGGCACA	97042
QY	45	ThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCysPheAspTyrVal	64
DB	97043	ACTGCATCCAAACATGACTTGGCGAGTCTTTTGTAGTGTCTGTCTTTGACTATGTG	97102
QY	65	LeuProProlleLeuGlnCysGlnSerGlyHisLeuValCysSerAsnCysArgProLys	84
DB	97103	TTGCCACCTATTCTTCAGTGTGAGAGTGGCCATCTTGTTTGTAGCACTGTGCCGCCAAA	97162
QY	85	LeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsnLeuAlaMetGlu	104
DB	97163	CTTACATGTTGTCACCTTGGCGGGGCCCATTTGGGATCCATTTGGCACTTGGCTATGGAG	97222
QY	105	LysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGlyCysGluLeuThr	124
DB	97223	AAAGTGGCCACTCAGTACTCTTCTTGTAAATATGCTCTTCTGGATGTGAATAACT	97282
QY	125	LeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArgProTyrSerCys	144
DB	97283	CTGCCACACACCCAAAGAGCAGACAGAGAGCTCTGTGAGTTCAGGCCTTACTCTGTC	97342
QY	145	ProCysProGlyAlaSerCysLysTrpGlnGlySerLeuAspAlaValMetProHisLeu	164
DB	97343	CCCTGCCCTTGGTCTTCTGTAAAGTGGCAAGGCTCCTTGGATGCCCTCATGCCCCACCTG	97402
QY	165	MetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleValPheLeuAlaThr	184
DB	97403	ATGCATCAGCACAAGTCCATTACCACCTGCAAGGAGAGATATAGTTTTCTTGTCTACA	97462
QY	185	AspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSerCysPheGlyPheHis	204
DB	97463	GACATTAACCTTCTTGGTCTGTGTGACTGGGTGATGATGACAGTCTTGTGTTTGGCTTTCAT	97522
QY	205	PheMetLeuValLeuGlnLysGlnGluLysTyrAspGlyHisGlnGlnPhePheAlaIle	224
DB	97523	TTTCATGTTAGTCTTGGAGAAACAAGAAAAATATGATGGTCATCAGCAGTCTTTTGCATTT	97582
QY	225	ValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArgLeuGluLeuAsn	244
DB	97583	GTACAACTGATAGGAACAAGCAAGCAAGCTGAAAAATTTTGCATATCGACTTGAAGTAAAT	97642
QY	245	GlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHisGlnGlyIleAla	264
DB	97643	GGTCATAGCGCGGCAATTGACTTGGGAAGGAGCTCTCGGTCTATTTCATGAGGAATTGCA	97702
QY	265	ThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAlaGlnLeuPheAla	284
DB	97703	ACAGCCATTATGAATAGTGTGCTGCTTGTAGTGTGTTTGACACAGCAGATTGACAGCTTTTGCA	97762

Qy 285 GluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298  
Db 97763 GAAATGGCAATTAGGCATCAATGTAACTATTTCATGTGT 97804

Search completed: April 25, 2005, 08:26:52  
Job time : 5084 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 25, 2005, 06:39:19 ; Search time 614 Seconds  
(without alignments)  
2873.099 Million cell updates/sec

Title: US-10-679-246-2

Perfect score: 1611

Sequence: 1 MWIIIFLPYVFISEMSRQ.....IAQLFAENGLGINTISMC 298

Scoring table:

BLSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framed\_p2n.model -DEV=xlh  
-O=/cn2.1/USPTO.spool/US10679246/runat\_22042005.185510.7675/app\_query.fasta\_1.455  
-DB=N\_Geneseq\_16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=15 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10679246 @CN 1.1 470 @runat\_22042005.185510.7675 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001as:\*

5: Geneseq2001bs:\*

6: Geneseq2002as:\*

7: Geneseq2002bs:\*

8: Geneseq2003as:\*

9: Geneseq2003bs:\*

10: Geneseq2003cs:\*

11: Geneseq2003ds:\*

12: Geneseq2004as:\*

13: Geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	1274	4 AAC67281	AAC67281 Human Siah
2	1611	100.0	2128	13 ADS34361	ADS34361 POSH prot
3	1611	100.0	2440	11 ADI31262	ADI31262 Human cdn
4	1611	100.0	2924	5 ABV28804	ABV28804 Human pro
5	1611	100.0	2924	5 ABV25468	ABV25468 Human pro

6	1611	100.0	2924	5	ABV22970	Human pro
7	1611	100.0	6107	4	AAL05382	Human rep
8	1611	100.0	6107	4	ABL98246	Human tes
9	1603	99.5	2829	11	ADM01584	Human cdn
10	1603	99.5	2829	13	ADM01584	Human cdn
11	1565	97.1	4090	10	ADF81981	Leukaemia
12	1540	95.6	1540	13	ADS34365	POSH prot
13	1540	95.6	2048	13	ADS34360	POSH prot
14	1537	95.4	2454	13	ADS34359	POSH prot
15	1536	95.3	1968	6	ABI99429	Mouse isc
16	1532	95.1	1886	13	ADS34363	POSH prot
17	1527	94.8	2034	13	ADS34364	POSH prot
18	1489	92.4	1884	2	AA764820	Tumour su
19	1334.5	82.8	3128	10	ADI40343	Human pur
20	1206	74.9	2502	8	ACC50293	Breast ca
21	1199.5	74.5	2198	6	ABI99430	Mouse isc
22	1199	74.4	975	13	ADR25256	Breast ca
23	1198	74.4	2240	8	ACC50292	Breast ca
24	1198	74.4	2240	12	ADP13401	Renal cel
25	1198	74.4	2240	13	ADR14122	Human NF-
26	1198	74.4	2240	13	ADP54830	Human PRO
27	1153.5	71.6	4647	4	ABL15923	Drosophil
28	1153.5	71.6	8418	4	ABL15922	Drosophil
29	1153.5	71.6	12029	4	ABL15508	Drosophil
30	702.5	43.6	3349	4	ABL21784	Drosophil
31	530	32.9	1226	4	ABL21785	Drosophil
32	514.5	31.9	2363	4	ABL21782	Drosophil
33	485.5	30.1	388	2	AAQ60440	Human bra
34	474	29.4	1535	3	AAC98856	Human pan
35	447	27.7	1420	3	AAZ90582	Maize SIN
36	446	27.7	1428	3	AAC43673	Zea mays
37	446	27.7	360	8	ABX44846	Bovine ES
38	434.5	27.0	1599	3	AAC51542	Arabidops
39	434.5	27.0	1601	3	AAC38735	Arabidops
40	429	26.6	984	3	AAC43003	Arabidops
41	429	26.6	984	6	ABZ12658	Arabidops
42	426	26.4	257	10	ACAS5799	Human sig
43	426	26.4	257	12	ADI55595	Human pol
44	425	26.4	1080	10	ADB78903	Rice tran
45	423	26.3	1264	3	AAC36352	Arabidops

ALIGNMENTS

RESULT 1  
AAC67281  
ID AAC67281 standard; cdna; 1274 BP.  
XX  
AC AAC67281;  
XX  
DT 09-APR-2001 (first entry)  
XX  
DE Human Siah-lalpa coding sequence SEQ ID NO: 1.  
XX

Human; protein degradation; siah-mediated degradation protein; SNMP;  
SCF-complex protein; SCP; siah-lalpa; siah-1 interacting protein; SIP;  
Kw Skp1-associated F-box protein; SAF-1; SAF-2; SAD; cancer; cell division;  
Kw Skp1-associated destruction-box protein; inflammatory disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200077207-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 09-JUN-2000; 2000WO-US015873.  
XX  
PR 11-JUN-1999; 99US-00330517.  
XX  
(BURN-) BURNHAM INST.  
PI Reed JC, Matsuzawa S;  
XX

DR WPI: 2001-071273/08.  
 DR P-PSDB; AAB35157.  
 XX Siah-Mediated Degradation Protein, useful for drug screening, for  
 PT therapeutic applications and for functional genomics.  
 XX  
 PS  
 PS Claim 5; Page 95-97; 121pp; English.  
 XX  
 XX The present invention provides the protein and coding sequences of  
 CC several siah-mediated degradation proteins and SCF-complex proteins.  
 CC These are designated Siah-1alpha, Siah-1 interacting protein (SIP), which  
 CC encodes two proteins due to alternative splicing (SIP-L and SIP-S), Skp1-  
 CC associated F-box protein-alpha and beta and -2 (SAF-1alpha, SAF-1beta  
 CC and SAF-2) and Skp1-associated destruction-box protein (SAD). The  
 CC proteins and their coding sequences are useful in the diagnosis and  
 CC treatment of cancers, disorders where too little cell division occurs  
 CC such as bone marrow aplasia, immunodeficiencies and inflammatory  
 CC diseases including sepsis, fibrosis, arthritis and graft versus host  
 CC disease  
 CC  
 XX Sequence 1274 BP; 336 A; 273 C; 253 G; 412 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1,046-138 Length: 1274  
 Score: 1611.00 Matches: 238  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-679-246-2 (1-298) x AAC67281 (1-1274)  
 QY 1 MetValIleIleIlePheLeuLeuProTyrValPheIleSerGluMetSerArgGln 20  
 Db 274 ATGGTTAATAATTATTTCTCCGCTCTTATGATATTATTTTTCAGAAATGAGCGCTCAG 333  
 QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla 40  
 Db 334 ACTGCTACAGCAATACCTACCGGTACTTCAATGTCACCAATCCAGAGGGTGGCTGCC 393  
 QY 41 LeuThrGlyThrThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCys 60  
 Db 394 CTGACTGGCACTGATCCATCCACATGACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGC 453  
 QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80  
 Db 454 TTTGACTATGTGTACGGCCCATTTCTCAATGTCAGAGTGGCCATCTTGTGTAGCAAC 513  
 QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
 Db 514 TGTGCCCCAAGCTCACATGTTGTCCAACTTGGCGGGCCCTTGGGATCCATTCGCAAC 573  
 QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
 Db 574 TTGGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATATGCGTCTCTGGA 633  
 QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140  
 Db 634 TGTGAATAAATCTGCCCCACACACAGAAAGACAGACCATGAAGAGCTCTGTGAGTTAGG 693  
 QY 141 ProTyrSerCysProCysProGlyValaSerCysLysTyrGlnGlySerLeuAspAlaVal 160  
 Db 694 CCTATTTCCTGCTGGTCCCTGGTGGTCTTCTGTAAATGGCAAGGCTCTCTGGATGCTGA 753  
 QY 161 MetProHisLeuMetHisGlnHisGlySerIleThrThrLeuGlnGlyGluAspIleVal 180  
 Db 754 ATGCCCATCTGATGCATCAGCAATAAGTCCATTACACCCCTACAGGAGGATATAGTT 813  
 QY 181 PheLeuAlaThrAspIleAsnLeuProGlyValAlaValAspTyrValMetMetGlnSerCys 200  
 Db 814 TTTCTTGCTACAGACATTAATCTTCCGTGGTCTGTGACTGGGTGATGATGACGCTCTGT 873  
 QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAspGlyHisGlnGln 220

Db 874 TTTGGCTTTTCACCTTCATGTAGTCTTAGAGAAACAGGAAAAATACGATGTCACCGACG 933  
 QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240  
 Db 934 TTTCTTGGCAATCGTACAGCTGTAGGACACACGCAAGCAAGCTGAAATTTTGTCTTACCGA 993  
 QY 241 LeuGluLeuAsnGlyHisArgArgGluLeuThrTyrGluAlaThrProArgSerIleHis 260  
 Db 994 CTTGAGCTAAATGCTCATAGCGACGATTCATTGGGAAGGACTCTCGATCTATTTCAT 1053  
 QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280  
 Db 1054 GAAGGAATTGCAACAGCCATTATGAAATAGCGACTGTCTAGTCTTTTGACACCGCATTTGCA 1113  
 QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298  
 Db 1114 CAGCTTTTTCAGAAAAATGGCAATTTAGGCAATCAATGTAATTTTCAATGTGT 1167  
 RESULT 2  
 ADS34361/C  
 ID ADS34361 standard; DNA; 2128 BP.  
 XX  
 AC ADS34361;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE POSH protein associated DNA #115.  
 XX  
 KW ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;  
 KW anticonvulsant; antiviral; neuroleptic; central nervous system;  
 KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI;  
 KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;  
 KW anti-cancer agent; secretory pathway trafficking inhibitor;  
 KW neurological disorder progression disorder; Alzheimer's disease;  
 KW Parkinson's disease; Huntington's disease; schizophrenia;  
 KW Niemann-Pick's disease.  
 XX Homo sapiens.  
 XX WO2004078130-A2.  
 XX 16-SEP-2004.  
 XX  
 XX 02-MAR-2004; 2004WO-US006308.  
 XX  
 XX 03-MAR-2003; 2003US-0451437P.  
 XX 05-MAR-2003; 2003US-0452284P.  
 XX 19-MAR-2003; 2003US-0455760P.  
 XX 20-MAR-2003; 2003US-0456640P.  
 XX 03-APR-2003; 2003US-0460526P.  
 XX 04-APR-2003; 2003US-0460792P.  
 XX 21-APR-2003; 2003US-0464285P.  
 XX 09-MAY-2003; 2003US-0469462P.  
 XX 15-MAY-2003; 2003US-0471378P.  
 XX 20-MAY-2003; 2003US-0472327P.  
 XX 30-MAY-2003; 2003US-0474706P.  
 XX 03-JUN-2003; 2003US-0475825P.  
 XX 17-JUN-2003; 2003US-0479317P.  
 XX 19-JUN-2003; 2003US-0480215P.  
 XX 19-JUN-2003; 2003US-0480376P.  
 XX 08-AUG-2003; 2003US-0493860P.  
 XX 28-AUG-2003; 2003US-0498634P.  
 XX 16-SEP-2003; 2003US-0503931P.  
 XX 10-NOV-2003; 2003WO-US035712.  
 XX 05-FEB-2004; 2004WO-US003600.  
 XX 02-MAR-2004; 2004US-0549896P.  
 XX  
 XX (PROT-) PROTEOLOGICS INC.  
 XX Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;  
 XX Greener T;  
 XX



CC detected in a non-diseased sample, where an altered level of the detected  
CC hybridisation complexes correlates with the presence of an  
CC immunopathological condition. Also disclosed are an expression profile  
CC comprising a microarray and a plurality of detectable complexes and a  
CC method for identifying a plurality of polynucleotide probes. The cDNAs  
CC are useful as hybridisable array elements in a microarray for monitoring  
CC the expression of target polynucleotides. The microarray can be used in  
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,  
CC ulcerative colitis, hyperesoinophilias, irritable bowel syndrome,  
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in  
CC identifying agents for the treatment of the diseases. The microarray may  
CC also be used in drug discovery and development, toxicological and  
CC carcinogenicity studies, forensics or pharmacogenomics. The composition  
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or  
CC genomic fragments. This sequence represents a human cDNA of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification but was obtained in electronic format directly  
CC from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 2440 BP; 685 A; 463 C; 472 G; 817 T; 0 U; 3 Other;

## Alignment Scores:

Pred. No.: 2,51e-138 Length: 2440  
Score: 1611.00 Matches: 298  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-10-679-246-2 (1-298) x AD131262 (1-2440)

QY 1 MetValIleIlePheLeuLeuProThrValPheIleSerGluMetSerArgGln 20  
DB 507 ATGTTTATTAATTAATTTCTCTGCTCTCTTAATTAATTTTTCAGAAATGAGCGTCTAG 566  
QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla 40  
DB 567 ACTGCTACAGCATTAACCTACCGGTACTCTGAAGTGTCCACCATCCAGAGGGTGCCTGCC 626  
QY 41 LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys 60  
DB 627 CTGACTGGCAACTGCATCCAAATGACTTGGCAGATCTTTTGGAGTCTTTTGGAGTCTG 686  
QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80  
DB 687 TTTGACTATGTGTACCGCCCATCTTCTCATGTCAGAGTGGCCATCTGTGTTGTAGCAAC 746  
QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
DB 747 TGTGCGCCAAAGCTACATGTTGTCCAACTTGCGGGGGCCCTTTGGGATCCATTCCGAAC 806  
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
DB 807 TTGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAATAATGCGTCTTTGGA 866  
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluLeuCysGluPheArg 140  
DB 867 TGTGAATTAATCTGTCACACACACAGAAAAGCAGACCATGAGAGCTCTGTGAGTTTAGG 926  
QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTrpGlnGlySerLeuAspAlaVal 160  
DB 927 CTTTATTCCTGTCGCGCCCTGGTGTCTCTGTAATGGAAGGCTCTCTGGATGCTGTA 986  
QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180  
DB 987 ATGCCCCATCTGATGCATCAGCATTAATGCTTACACACCTTACAGGGAGAGGATATAGTT 1046  
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSerCys 200  
DB 1047 TTTCTTGCTACAGACATTAATCTTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1106  
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAspGlyHisGlnGln 220

Db 1107 TTTGGCTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAAATACGATGTCACACGAG 1166  
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240  
Db 1167 TTTCTTCCAACTGTACAGCTGATAGGAACACGCAAGCAAGCTGAAATTTTGTCTTACCGA 1226  
QY 241 LeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260  
Db 1227 CTTGAGCTAAATGTCATAGCGCAGCATTCCTCTGGGAAGCGACTCTCGCATCTATTTCAT 1286  
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280  
Db 1287 GARGGAATTGCAACAGCCATTATGAATAGGACCTGTCTAGTCTTTGACACAGCATTTGCA 1346  
QY 281 GlnLeuPheAlaGluAsnGlyValLeuGlyLeuValThrIleSerMetCys 298  
Db 1347 CAGCTTTTGGAGAAATGGCAATTTAGGCATCAATGTAACTATTTCATGTGT 1400

## RESULT 4

ABV28804  
ID ABV28804 standard; cDNA; 2924 BP.  
XX AC ABV28804;  
XX DT 16-SEP-2002. (first entry)  
XX DE Human prostate expression marker cDNA 28795.  
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX KW Pharmacogenomic marker; gene; ss.  
XX OS Homo sapiens.  
XX PN WO200160860-A2.  
XX PD 23-AUG-2001.  
XX PF 20-FEB-2001; 2001WO-US005171.  
XX PR 17-FEB-2000; 2000US-0183319P.  
XX PR 16-MAR-2000; 2000US-0189862P.  
XX PR 25-MAY-2000; 2000US-0207454P.  
XX PR 09-JUN-2000; 2000US-0211314P.  
XX PR 18-JUL-2000; 2000US-0219007P.  
XX PR 13-DEC-2000; 2000US-0255281P.  
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX PI Schlegel R, Endege WO, Monahan JE;  
XX DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 6058; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
XX specification or its complement. (I) is useful for: (a) assessing whether  
XX a patient is afflicted with prostate cancer; (b) monitoring the  
XX progression of prostate cancer in a patient; (c) assessing the efficacy  
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing  
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;  
XX (e) selecting a composition for inhibiting prostate cancer in a patient;  
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
XX determining whether prostate cancer has metastasized in a patient; (h)  
XX assessing the aggressiveness or indolence of prostate cancer in a patient  
XX ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 2924 BP; 871 A; 562 C; 562 G; 925 T; 0 U; 4 Other;

## Alignment Scores:

Pred. No.: 3,228-138 Length: 2924  
 Score: 1611.00 Matches: 298  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-679-246-2 (1-298) x ABV28804 (1-2924)

QY 1 MetValIleIlePheLeuLeuProTyrValPheIleSerGluMetSerArgGln 20  
 DB 753 ATGGTTATATATTTTCTCTGCTCTTATGATTTATTTTCAGAAATGAGCGGTGAG 812  
 QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla 40  
 DB 813 ACTGCTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTGCC 872  
 QY 41 LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys 60  
 DB 873 CTGACTGGCACAACCTGCAATGATGCTGGCGAGTCTTTTGGAGTGTCCAGTCTGC 932  
 QY 61 PheAspTyrValLeuProThrLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80  
 DB 933 TTTGACTATGTTGACCGCCCATCTTCAATGTGAGTGGCACTCTGTTGTAGCAAC 992  
 QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
 DB 993 TGTGCCCAAGCTCACATGTTGTGCAACTGTCGGGGCCCTTTGGATCCATTGCCAAC 1052  
 QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
 DB 1053 TTGGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGA 1112  
 QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluLysCysGluPheArg 140  
 DB 1113 TGTGAATTAACCTGTCACACACAGAAAGACAGACCATGAAGAGCTCTGTGAGTTAGG 1172  
 QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal 160  
 DB 1173 CTTTATTCCTGTCCGTGCTGCTGCTCTCTTAAATGCAAGGCTCTCTGGATGCTGA 1232  
 QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180  
 DB 1233 ATGCCCCCATCTGATGCATCAGCATAAGTCCATTACAACCCCTACGGGAGGATATAGTT 1292  
 QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetMetGlnSerCys 200  
 DB 1293 TTTCTGTCTACAGACATTAACTTCTGTGCTGCTGTTGACTGGGTGATGATGAGTCTGT 1352  
 QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAspGlyHisGlnGln 220  
 DB 1353 TTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAATATACGATCCACGACG 1412  
 QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240  
 DB 1413 TTTCTTGGCAATCGTACGCTGATAGAAACACCAAGCAGCTGCAAAATTTTCTTACCGA 1472  
 QY 241 LeuGluLeuAsnGlyHisArgArgArgLeuThrTyrGluAlaThrProArgSerIleHis 260  
 DB 1473 CTTAGCTAAATGGTTCATAGGCGAGATTGATTTGGGAGGACCTCTCGATCTATTAT 1532  
 QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280  
 DB 1533 GAAGGAATGGCAACAGCATTATGATAGGACTGTCTAGTCTTTTGACACACGAGATTGCA 1592  
 QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298  
 DB 1593 CAGCTTTTTCAGAGAAATGGCAATTTAGGCATCAATGTAATCTATTTCATGTGT 1646

## RESULT 5

ABV25468

ID ABV25468 standard; cDNA; 2924 BP.

XX AC

ABV25468;

XX DT

16-SEP-2002 (first entry)

XX DE

Human prostate expression marker cDNA 25459.

XX KW

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; pharmacogenomic marker; gene; ss.

XX OS

Homo sapiens.

XX PN

WO200160860-A2.

XX PD

23-AUG-2001.

XX PF

20-FEB-2001; 2001WO-US005171.

XX PR

17-FEB-2000; 2000US-0183313P.

PR PR

16-MAR-2000; 2000US-0189862P.

PR PR

25-MAY-2000; 2000US-0207454P.

PR PR

09-JUN-2000; 2000US-0211314P.

PR PR

18-JUL-2000; 2000US-0219007P.

PR PR

13-DEC-2000; 2000US-0255281P.

XX PA

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI

Schlegel R, Endege WO, Monahan JB;

XX PX

WPI; 2001-662795/76.

XX PT

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX PS

Claim 1; Page 5054-5055; 11750pp; English.

XX CC

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ

Sequence 2924 BP; 871 A; 562 C; 562 G; 925 T; 0 U; 4 Other;

## Alignment Scores:

Pred. No.: 3,228-138 Length: 2924  
 Score: 1611.00 Matches: 298  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-679-246-2 (1-298) x ABV25468 (1-2924)

QY 1 MetValIleIlePheLeuLeuProTyrValPheIleSerGluMetSerArgGln 20

DB 753 ATGGTTATATATTTTCTCTGCTCTTATGATTTATTTTCAGAAATGAGCGGTGAG 812

QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla 40

DB 813 ACTGCTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTGCC 872

QY 41 LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys 60

DB 873 CTGACTGGCACAACATGCATCCAAATGACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGC 932

QY 61 PheAspTyrValLeuProPheLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80  
 DB 933 TTTGACTATGTTACCGCCCATCTTCAATGTCAGAGTGGCCATCTTGTGTAGCAAC 992  
 QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
 DB 993 TGTGCGCCCAAGCTCACATGTTGTCCAACTTGTCCGGGGCCCTTTGGGATCCATTGCGCAAC 1052  
 QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
 DB 1053 TTGGCTATGGAGAAAGTGCTTAATTCAGTACTTTCCCTGTAAATATGCGTCTCTGGA 1112  
 QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140  
 DB 1113 TGTGAATAAATCTGCGCACACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1172  
 QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal 160  
 DB 1173 CCTTATCTCTGCGTGGCTGCTTCTCTGTAATGGAAGGCTCTCTGGATGCTGTA 1232  
 QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180  
 DB 1233 ATGCCCATCTGATGATCAGATATGCTTACATCAACCTTACAGGAGAGAGATAGTT 1292  
 QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetMetGlnSerCys 200  
 DB 1293 TTTCTTGCTACAGACATTAATCTCTCGTGTCTGTTGACTGGGTGATGATGAGTCTGT 1352  
 QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAspGlyHisGlnGln 220  
 DB 1353 TTTGGCTTTCATCTTCAATGTTAGTCTTAGAGAAACAGAGAAATACGATGTCACGAGCAG 1412  
 QY 221 PhePheAlaIleValGlnLeuLeuGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240  
 DB 1413 TTCTTCCGAATGTCAGCTGATAGGACACAGCAAGAGCTGAAATTTTGTCTTACCGA 1472  
 QY 241 LeuGluLeuAsnGlyHisArgArgLeuThrThrGluAlaThrProArgSerIleHis 260  
 DB 1473 CTTGAGCTAAATGCTCATAGCGACGATGACTTTGGGAGGAGTCTCTCACTATTCAT 1532  
 QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280  
 DB 1533 GAGGAATTCACAGCATTATGATAGGAGCTGCTAGTCTTTGACACAGCATTGCA 1592  
 QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298  
 DB 1593 CAGCTTTTTCAGAAAAATGGCAATTTAGGCATCAATGTAATTTCCATGTGT 1646

RESULT 6  
 ID ABV22970 standard; cDNA; 2924 BP.  
 AC ABV22970;  
 XX 13-SEP-2002 (first entry)  
 DE Human prostate expression marker cDNA 22961.  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 OS Homo sapiens.  
 XX WO200160860-A2.  
 XX 23-AUG-2001.  
 XX 20-FEB-2001; 2001WO-US005171.  
 XX 17-FEB-2000; 2000US-0183319P.  
 XX 16-MAR-2000; 2000US-0189862P.  
 XX 25-MAY-2000; 2000US-0207454P.

09-JUN-2000; 2000US-0211314P.  
 18-JUL-2000; 2000US-0219007P.  
 13-DEC-2000; 2000US-0255281P.  
 (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 Schlegel R, Endege WO, Monahan JE;  
 WPI; 2001-662795/76.  
 Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.  
 Claim 1; Page 4080-4081; 11750pp; English.  
 The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV22213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing whether prostate cancer has metastasized in a patient; (g) determining the aggressiveness or indolence of prostate cancer in a patient assessing the aggressiveness or indolence of prostate cancer in a patient; (h) (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 SQ Sequence 2924 BP; 871 A; 562 C; 562 G; 925 T; 0 U; 4 Other;  
 Alignment Scores:  
 Pred. No.: 3.22e-138 Length: 2924  
 Score: 1611.00 Matches: 298  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0  
 US-10-679-246-2 (1-298) x ABV22970 (1-2924)  
 QY 1 MetValIleIleLeuPheLeuLeuProTyrValPheIleSerGluMetSerArgGln 20  
 DB 753 ATGGTTATAATTAATTTTCTCTGCTCTTATGTTATTTTTCAGAAATGAGCGGTGAG 812  
 QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSerGlnArgValProAla 40  
 DB 813 ACTGCTACAGCATTACCTACCGGTACCTCGAAGTGTCCACCATCCAGAGGGTGTCTGCC 872  
 QY 41 LeuThrGlyThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCys 60  
 DB 873 CTGACTGGCACAACTGCATCCCAATGACTTGGCGAGTCTTTTGAGTGTCCAGTCTGC 932  
 QY 61 PheAspTyrValLeuProPheLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80  
 DB 933 TTTGACTATGTTACCGCCCATCTTCAATGTCAGAGTGGCCATCTTCTTTGTAGCAAC 992  
 QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
 DB 993 TGTGCGCCCAAGCTCACATGTTGTCCAACTTGTCCGGGGCCCTTTGGGATCCATTGCGCAAC 1052  
 QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
 DB 1053 TTGGCTATGGAGAAAGTGCTTAATTCAGTACTTTCCCTGTAAATATGCGTCTCTGGA 1112  
 QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140  
 DB 1113 TGTGAATAAATCTGCGCACACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1172  
 QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal 160  
 DB 1173 CCTTATCTCTGCGTGGCTGCTTCTCTGTAATGGAAGGCTCTCTGGATGCTGTA 1232

QY 161 MetProHisMetHisGlnHisIleSerIleThrThriLeuGlnGlyGluAspIleVal 180  
Db 1233 ATCCCCCATCTGATCATCAGCAATAAGTCCATTACAAACCTACAGGAGAGGATATAGTT 1292  
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTirPValMetMetGlnSerCys 200  
Db 1293 TTTCCTGCTACACACATTAACTCTCTGGTGTCTTTGACTGGGTGATGATGCAGTCTCTGT 1352  
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAspGlyHisGlnGln 220  
Db 1353 TTTCGGCTTTCACCTTCATGTTAGTCTTAGAGAAACAGCAAAATACGATGTCACACGAG 1412  
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240  
Db 1413 TTCTTCGCATCATCTACAGCTGATAGCAACGCAAGCAAGCTGAAAAATTTTGTCTTACCGA 1472  
QY 241 LeuGluLeuAsnGlyHisArgArgLeuThrTirPLeuAlaThrProArgSerIleHis 260  
Db 1473 CTTCGACTAAATGCTCATAGGCGACGATTGACTTGGAGGCGACTCTCGATCTATTCAT 1532  
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280  
Db 1533 GAAGGAATTGCAACGCCATTATGAATAGCGACTGTAGTCTTTGCACACGACGATTGCA 1592  
QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298  
Db 1593 CAGCTTTTTCAGAAAAATGCAATTTAGGCATCAATGTAACCTATTTCCATGTGT 1646  
RESULT 7  
AAL05382  
ID AAL05382 standard; DNA; 6107 BP.  
XX AAL05382;  
AC AAL05382;  
DT 21-NOV-2001 (first entry)  
DE Human reproductive system related antigen DNA SEQ ID NO: 8070.  
XX Human reproductive system related antigen; reproductive system disorder;  
KW Human; reproductive system related antigen; cancer; gene therapy; ds.  
XX Homo sapiens.  
XX WO200155320-A2.  
XX 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US001339.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214686P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216980P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0232081P.  
PR 14-SEP-2000; 2000US-0233977P.  
PR 14-SEP-2000; 2000US-0233988P.  
PR 14-SEP-2000; 2000US-0233999P.  
PR 14-SEP-2000; 2000US-0234000P.  
PR 14-SEP-2000; 2000US-0234011P.  
PR 14-SEP-2000; 2000US-0234063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.



08-NOV-2000; 2000US-0246527P.  
 08-NOV-2000; 2000US-0246528P.  
 08-NOV-2000; 2000US-0246532P.  
 08-NOV-2000; 2000US-0246609P.  
 08-NOV-2000; 2000US-0246610P.  
 08-NOV-2000; 2000US-0246611P.  
 08-NOV-2000; 2000US-0246613P.  
 17-NOV-2000; 2000US-0249207P.  
 17-NOV-2000; 2000US-0249208P.  
 17-NOV-2000; 2000US-0249209P.  
 17-NOV-2000; 2000US-0249210P.  
 17-NOV-2000; 2000US-0249211P.  
 17-NOV-2000; 2000US-0249212P.  
 17-NOV-2000; 2000US-0249213P.  
 17-NOV-2000; 2000US-0249214P.  
 17-NOV-2000; 2000US-0249215P.  
 17-NOV-2000; 2000US-0249216P.  
 17-NOV-2000; 2000US-0249217P.  
 17-NOV-2000; 2000US-0249218P.  
 17-NOV-2000; 2000US-0249244P.  
 17-NOV-2000; 2000US-0249245P.  
 17-NOV-2000; 2000US-0249264P.  
 17-NOV-2000; 2000US-0249265P.  
 17-NOV-2000; 2000US-0249266P.  
 17-NOV-2000; 2000US-0249297P.  
 17-NOV-2000; 2000US-0249299P.  
 17-NOV-2000; 2000US-0249300P.  
 01-DEC-2000; 2000US-0250160P.  
 01-DEC-2000; 2000US-0250391P.  
 05-DEC-2000; 2000US-0251030P.  
 05-DEC-2000; 2000US-0251988P.  
 05-DEC-2000; 2000US-0256719P.  
 08-DEC-2000; 2000US-0251479P.  
 08-DEC-2000; 2000US-0251856P.  
 08-DEC-2000; 2000US-0251868P.  
 08-DEC-2000; 2000US-0251869P.  
 08-DEC-2000; 2000US-0251989P.  
 08-DEC-2000; 2000US-0251990P.  
 11-DEC-2000; 2000US-0254097P.  
 03-JAN-2001; 2001US-0259678P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Rosen CA, Barash SC, Ruben SM;  
 WPI; 2001-465570/50.  
 Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition.  
 Disclosure; SEQ ID NO 8070; 1297pp + Sequence Listing; English.  
 The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a genomic sequence encoding a protein of the invention  
 Alignment Scores:  
 Pred. No.: 8,79e-138 Length: 6107  
 Score: 1611.00 Matches: 298  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-679-246-2 (1-298) x AAL05382 (1-6107)  
 Qy 1 MetValIleIleIlePheLeuLeuProTyrValPheIleSerGluMetSerArgGln 20  
 Db 3437 ATGGTTATATATATTTCTCTGCTGCTTATGTAATTTTTCAGAAATGAGCGGTGAG 3496

Qy 21 ThrAlaThrAlaLeuProThrGlyThrSerIysCysProProSerGlnArgValProAla 40  
 Db 3497 ACTGCTACAGCAATTACCTACCGGTACCTCGAAGTGTCCACCATCCAGAGGGTCCCTGCC 3556  
 Qy 41 LeuThrGlyThrThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCys 60  
 Db 3557 CTGACTGGCACAACCTGCATCCACAATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGC 3616  
 Qy 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80  
 Db 3617 TTTGACTAATGTTTACCGCCCATCTTCAATGTACAGAGTGGCCATCTTTGTTAGCAAC 3676  
 Qy 81 CysArgProIysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
 Db 3677 TGTGCCCAAGCTCACATGTTGTCCAATTTGCCGGGGCCCTTTGGGATCCATCTCGCAAC 3736  
 Qy 101 LeuAlaMetGluIysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
 Db 3737 TTGGCTATGGAGAAAGTGGCTAATTTCAATGTTTCCCTGTAAATATGCGTCTTCTGGA 3796  
 Qy 121 CysGluIleThrLeuProHisThrGluIysAlaAspHisGluLeuLeuCysGluPheArg 140  
 Db 3797 TGTGAATAACTCTGCCACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTAGG 3856  
 Qy 141 ProTyrSerCysProCysProGlyAlaSerCysIysTyrGlnGlySerLeuAspAlaVal 160  
 Db 3857 CATTATCTCTCGCGTCCCTGCTTCTTAAATGGCAAGGCTCTCTGGATGCTGTA 3916  
 Qy 161 MetProHisLeuMetHisGlnHisIysSerIleThrThrLeuGlnGlyGluAspIleVal 180  
 Db 3917 ATGCCCCATCTGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3976  
 Qy 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetMetGlnSerCys 200  
 Db 3977 TTTCTTGTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4036  
 Qy 201 PheGlyPheHisPheMetLeuValLeuGluIysGlnGlyIysTyrAspGlyHisGlnGln 220  
 Db 4037 TTTGGCTTTCACTTTCATGTTAGTCTTAGAGAAACAGGAAATATAGTGTGTCACCGAG 4096  
 Qy 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240  
 Db 4097 TTTCTTGGCAATGTTACAGCTGATAGGACAGCAAGCAAGCTGAAATTTTCTTACCGA 4156  
 Qy 241 LeuGluLeuAsnGlyHisArgArgLeuThrThrGluAlaThrProArgSerIleHis 260  
 Db 4157 CTTGAGCTAAATGGTATAGGCGAGATTGACTTGGGAAGCGACTCTCTGATCTATTAT 4216  
 Qy 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280  
 Db 4217 GAAGGAATTCACACAGCCATTAATGATAGCGACTGTCTAGTCTTTGACACACAGCATTGCA 4276  
 Qy 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298  
 Db 4277 CAGCTTTTTCAGAAATGCGCAATTTAGGCATCAATGTAACTATTATTTCCATGTGT 4330  
 RESULT 8  
 ABL98246  
 ID ABL98246 standard; DNA; 6107 BP.  
 XX  
 AC ABL98246;  
 XX  
 DT 21-JUN-2002 (first entry)  
 XX  
 DE Human testicular antigen encoding DNA fragment SEQ ID NO: 2898.  
 XX  
 KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
 KW reproductive system disorder; urinary system disorder; gene therapy;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disease; infection; cytostatic; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX

PN WO200155317-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US001329.  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234224P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 02-OCT-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
PI

Db	4097	TTCTTCGCAATCGTACAGCTGATAGGAACACGACGACGCTGAAATTTTGTACCGA	4155
Qy	241	LeuGluleuAsnGlyHisArgArgGLeuThrTIPgluAlaThrProArgSerIleHis	260
Db	4157	CTTGAGCTAAATGGTCATAGGCGACGATTGACITGGGAAGCGACTCTCGATCTATTTCAT	4216
Qy	261	GlulGlyleAlaThrAlaIleMetAsnSerAspCysIleuValPheAspThrSerIleAla	280
Db	4217	GAAGGAATTCGCAACAGCCATTATCAATAGCGACTGTCTAGTCTTTGACACGACGATTGCA	4276
Qy	281	GlnIleuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys	298
Db	4277	CAGCTTTTTCGAGAAATGCGCAATTTAGGCATCAATGTAACTATTTCATGTGT	4330
RESULT 9			
ID	ADM01584	standard; cDNA; 2829 BP.	
XX	XX		
XX	ADM01584;		
XX	XX		
DT	20-MAY-2004	(first entry)	
XX	XX		
DE	Human cDNA of the invention SEQ ID NO:269.		
XX	ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.		
KW	XX		
OS	Homo sapiens.		
XX	XX		
PN	EPI347046-A1.		
XX	XX		
PD	24-SEP-2003.		
XX	XX		
PF	12-APR-2002; 2002EP-00008400.		
XX	XX		
PR	22-MAR-2002; 2002JP-00137785.		
XX	XX		
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.		
XX	XX		
PI	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;		
PI	Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;		
PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;		
XX	XX		
DR	WPI; 2003-723558/69.		
DR	P-PSDB; ADM04027.		
XX	New polynucleotides and polypeptides are useful in gene therapy, for		
PT	developing a diagnostic marker or medicines for regulating their		
PT	expression and activity, or as a target of gene therapy.		
PS	Claim 1; SEQ ID NO 269; 305pp; English.		
XX	The invention relates to a novel human polynucleotide and the encoded		
CC	polypeptide. A polynucleotide of the invention may have a use in gene		
CC	therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful		
CC	as a primer for synthesizing the polynucleotide or as a probe for		
CC	detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are		
CC	useful in gene therapy, for developing a diagnostic marker or medicines		
CC	for regulating their expression and activity, or as a target of gene		
CC	therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides		
CC	are useful as pharmaceutical agents. The present sequence represents a		
CC	cDNA sequence of the invention.		
XX	Sequence 2829 BP; 756 A; 566 C; 553 G; 954 T; 0 U; 0 Other;		
SQ			
Alignment Scores:			
Pred. No.:	1.68e-137	Length:	2829
Score:	1603.00	Matches:	297
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatches:	1
Query Match:	99.50%	Indels:	0
DB:	11	Gaps:	0



Score:	1603.00	Matches:	297
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatches:	1
Query Match:	99.50%	Indels:	0
DB:	13	Gaps:	0
US-10-679-246-2 (1-298) x ADS34362 (1-2829)			
QY	1	MetValleillePheLeuLeuProProTyrValPheIleSerGluMetSerArgGln	20
DB	1407	ATGGTTATAATTATTTTCTCTGCTCTCTATGTTATTTATTTACAGAAATGAGCGGTGAG	1466
QY	21	ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla	40
DB	1467	ACTGCTACAGCATTAACCTACCGGTACCTCGAAGTGTCCACCATCCAGAGGTGCTGCC	1526
QY	41	LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys	60
DB	1527	CTGACTGGCACAACATGCATCCACAATGACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGC	1586
QY	61	PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn	80
DB	1587	TTTGACTATGTGTACCGGCCATCTTCAATGTTCAGAGTGGCCATCTTTTGTAGCAAC	1646
QY	81	CysArgProIysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn	100
DB	1647	TGTGCCCAAGCTCATGTGTGTCCAACTTGGCGGGCCCTTTGGGATCCATTCGCAAC	1706
QY	101	LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly	120
DB	1707	TTGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAATATATCGTCTTCGGA	1766
QY	121	CysGluThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg	140
DB	1767	TGTGAATAACTCTGCCACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTAGG	1826
QY	141	ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal	160
DB	1827	CCTTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1886
QY	161	MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal	180
DB	1887	ATGCCCATCTGATGCATCAGCATAGTCCATTAACACCTACAGGGAGGATATAGTT	1946
QY	181	PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetMetGlnSerCys	200
DB	1947	TTTCTTGTCTACAGCAATTAATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2006
QY	201	PheGlyPheHisPheMetLeuValLeuGluLysGlnGlnLysTyrAspGlyHisGlnGln	220
DB	2007	TTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAATACGATGTCACACGAG	2066
QY	221	PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg	240
DB	2067	TTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAAATTTTGTCTACCGA	2126
QY	241	LeuGluLeuAsnGlyHisArgArgArgLeuThrTyrGluAlaThrProArgSerIleHis	260
DB	2127	CTTGAGCTAAATGGTCTAGCGGACGATGTGACTTGGGAGGAGCATCTCTCATTTATTCAT	2186
QY	261	GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla	280
DB	2187	GAAGGAATTGCAACAGCCATTTATGATAGCGACTGTCTAGTCTTTGCCACGACATTGCA	2246
QY	281	GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys	298
DB	2247	CAGCTTTTTCAGAAAAATGGCAATTTAGGCATCAATGTAACTATTTTCCATGTGT	2300
RESULT 11			
ADF81981/c			
ID	ADF81981 standard; DNA; 4090 BP.		
XX			
AC	ADF81981;		
XX	26-FEB-2004	(first entry)	
DE	Leukaemia-related DNA sequence #2537.		
KW	Cytostatic; Gene therapy; leukaemia; ss.		
OS	Unidentified.		
FN	WO2003039443-A2.		
PD	15-MAY-2003.		
PP	04-NOV-2002; 2002WO-EP012303.		
PR	05-NOV-2001; 2001EP-00126244.		
PR	30-APR-2002; 2002EP-00009758.		
XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.		
PA	(UYLU-) UNIV LUDWIG MAXIMILIANS.		
PA	(HAPE/) HAERLACH T.		
PA	(SCHO/) SCHOCH C.		
PA	(KERN/) KERN W.		
XX	Haerlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;		
PI	Eils R, Broers B, Mergenthaler S;		
XX	WPI; 2003-505037/47.		
XX	Determining the subtype of leukemia cells and whether a patient sample		
PT	contains leukemia cells or other cells, useful for treating leukemia, a		
PT	comprises determining the expression profile of a group of markers in a		
PT	patient sample.		
XX	Disclosure; SEQ ID NO 2537; 2938pp; English.		
CC	The present invention relates to a method (M1) for determining the		
CC	subtype of leukaemia cells and whether a patient sample contains		
CC	leukaemia cells. The method comprises determining the expression profile		
CC	of a group of markers in a patient sample. The method is useful for		
CC	determining the presence of leukaemia cells, its types or subtypes, and		
CC	for the preparation of a medicament for treating leukaemia.		
XX	Sequence 4090 BP; 1233 A; 840 C; 768 G; 1189 T; 0 U; 60 Other;		
Alignment Scores:			
Pred. No.:	8.89e-134	Length:	4090
Score:	1565.00	Matches:	295
Percent Similarity:	98.33%	Conservative:	0
Best Local Similarity:	98.33%	Mismatches:	3
Query Match:	97.14%	Indels:	2
DB:	10	Gaps:	0
US-10-679-246-2 (1-298) x ADF81981 (1-4090)			
QY	1	MetValleillePheLeuLeuProProTyrValPheIleSerGluMetSerArgGln	20
DB	3338	ATGGTTATAATTATTTTCTCTGCTCTCTATGTTATTTATTTACAGAAATGAGCGGTGAG	3279
QY	21	ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla	40
DB	3278	ACTGCTACAGCATTAACCTACCGGTACCTCGAAGTGTCCACCATCCAGAGGNTGCTGCC	3219
QY	41	LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys	60
DB	3218	CTGACTGGCACAACATGCATCCACAATGACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGC	3159
QY	61	PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn	80
DB	3158	TTTGACTATGTGTACCGCCCATCTTCAATGTACAGTGGCCATCTTTTGTAGCAAC	3099
QY	81	CysArgProIysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn	100

```
Db 3098 TGTGCGCCAAAGCTCACAATGTTGTCACATCTTGTCCAACTTGTCCGGGGGCCCTTTGGGATCCATTCGCAAC 3039
Qy 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120
Db 3038 TTGGCTATGAGAAAGTGCGTAATTCAGTACTTTTCCCTCTGTAATAATGCGTCTTCTGGA 2979
Qy 121 CysGluLeuThrLeuProHisThrGluLysAlaAspHisGluGluLeuLysGluPheA 140
Db 2978 TGTGAATAAATCTGNCACACACAGAAAAANGCAGACCATGAANANGCTCTGTGAGCTTTA 2919
Qy 140 igProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlav 160
Db 2918 GGCCTTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2859
Qy 160 alMetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleV 180
Db 2858 TAATGCCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2799
Qy 180 alPheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetGlnSerC 200
Db 2798 TTTTTCCTTCTGATCAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2739
Qy 200 YspPheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAspGlyHisGlnG 220
Db 2738 GTTTTGGCTTTTCACTTCATGTTAGTCTTTAGAGAAACAGGAAATAACGATGGTCACGAC 2679
Qy 220 lnPhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrA 240
Db 2678 AGTTCTTCGCAATCGTACAGCTGATAGGACACGCAAGCAAGCTGAAATTTTGTCTTACC 2619
Qy 240 rgLeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleH 260
Db 2618 GACTTGAGCTAAATGCTCATAGCGCAGATTGACTTGGGAAGCGACTCTCTCGATCTATTTC 2559
Qy 260 isGluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleA 280
Db 2558 ATGAAGGAATTGCAACAGCATTTATGAATAGGCACTGTCTAGTCTTTTGACACAGCATTTG 2499
Qy 280 laGlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298
Db 2498 CACAGCTTTTTCAGAAATGCAATTTAGGCATCAATGTAATTTTCAATGTTGT 2443
RESULT 12
ADS34365
ID ADS34365 standard; DNA; 1540 BP.
AC ADS34365;
XX
XX 02-DEC-2004 (first entry)
XX
XX POSH protein associated DNA #119.
KW ds; gene; cytostatic; nootropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; antiviral; neuroleptic; central nervous system;
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUD1;
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;
KW anti-cancer agent; secretory pathway trafficking inhibitor;
KW neurological disorder progression disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia;
KW Niemann-Pick's disease.
XX
OS Homo sapiens.
XX
XX WO2004078130-A2.
XX
XX 16-SEP-2004.
XX
XX 02-MAR-2004; 2004WO-US0006308.
XX
XX 03-MAR-2003; 2003US-0451437P.
XX 05-MAR-2003; 2003US-0452284P.
XX 19-MAR-2003; 2003US-0455760P.
XX 20-MAR-2003; 2003US-0456640P.
```

```
PR 03-APR-2003; 2003US-0460526P.
PR 04-APR-2003; 2003US-0460792P.
PR 21-APR-2003; 2003US-0464285P.
PR 09-MAY-2003; 2003US-0469462P.
PR 15-MAY-2003; 2003US-0471378P.
PR 20-MAY-2003; 2003US-0472327P.
PR 30-MAY-2003; 2003US-0474706P.
PR 03-JUN-2003; 2003US-0475825P.
PR 17-JUN-2003; 2003US-0479317P.
PR 19-JUN-2003; 2003US-0480215P.
PR 19-JUN-2003; 2003US-0480376P.
PR 08-AUG-2003; 2003US-0493860P.
PR 28-AUG-2003; 2003US-0498634P.
PR 16-SEP-2003; 2003US-0503931P.
PR 10-NOV-2003; 2003WO-US035712.
PR 05-FEB-2004; 2004WO-US003600.
PR 02-MAR-2004; 2004US-0549896P.
XX
XX (PROT-) PROTEOLOGICS INC.
PA Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
PI Greener T;
XX WPI; 2004-662346/64.
XX
PT Isolated, purified or recombinant complex, useful for identifying an
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
PT POSH-associated protein (POSH-AP).
XX
XX Disclosure; SEQ ID NO 129; 374pp; English.
XX
XX The invention relates to an isolated, purified or recombinant complex (I)
XX comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
XX or HERPUD1 and a ubiquitin ligase (b). Methods using (I). (a) or (b) are
XX useful for identifying an agent that modulates an activity of a POSH
XX polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
XX apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
XX of a protein through the secretory pathway, an agent that inhibits the
XX progression of a neurological disorder, an agent that modulates a POSH
XX function, an agent that modulates a HERPUD1 function. The methods can be
XX used for treating a viral infection, for inhibiting an activity of a POSH
XX -AP in a cell, for treating a POSH-associated disease in a subject. The
XX POSH-associated disease is viral infection, POSH-associated cancer or
XX POSH-associated neurological disorder. The methods are useful for
XX treating or preventing POSH-associated neurological disorder in a subject
XX e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
XX schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
XX nucleic acid of the invention.
SQ Sequence 1540 BP; 428 A; 313 C; 352 G; 447 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 4,74e-132 Length: 1540
Score: 1540.00 Matches: 285
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.28% Mismatches: 0
Query Match: 95.59% Indels: 4
DB: 13 Gaps: 1
US-10-679-246-2 (1-298) x ADS34365 (1-1540)
Qy 9 ProProTyrValPheIleSerGluMetSerArgGlnThrAlaThrAlaLeuProThrGly 28
Db 106 CGCCCC-----ACAGAAATGAGCCGTCAGACTGCTACAGCATTTACTACCGGT 153
Qy 29 ThrSerLysCysProProSerGlnArgValProAlaLeuThrGlyThrThrAlaSerAsn.48
Db 154 ACCTCGAAGTGTCCACCATCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 213
Qy 49 AsnAspLeuAlaSerLeuPheGluCysProValCysPheAspTyrValLeuProProfile 68
Db 214 AATGACTGGCGAGTCTTTTGTGAGTGTCCAGTGTCCAGTGTCTGTTGACTATGTGTGTACCGCCATT 273
```





```

Db      223 ACCTCGAAGTGTCCACCATCCAGAGGGTCCCTGCCCTGACTGGCACAACATGATCAAC 282
Qy      49  AsnAspLeuAlaSerLeuPheGluCysProValCysPheAspTyrValLeuProPhe 68
      |||||
Db      283 AATGACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTGACTATGTGTACCGCCCAT 342
Qy      69  LeuGlnCysGlnSerGlyHisLeuValCysSerAsnCysArgProLysLeuThrCysCys 88
      |||||
Db      343 CTTCAATGTCCAGAGTGGCCATCTTGTGTAGCAACTGTGCCCAAGCTCAATGTGT 402
Qy      89  ProThrCysArgGlyProLeuGlySerIleArgAsnLeuAlaMetGluLysValAlaAsn 108
      |||||
Db      403 CCAACTTGGCGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGGAGAAGTGGCTAAT 462
Qy      109 SerValLeuPheProCysLysTyrAlaSerSerGlyCysGluIleThrLeuProHisThr 128
      |||||
Db      463 TCAGTACTTTTCCCTGTAATATGCGTCTCTGGATGTGAATAACTCTGCCACACACA 522
Qy      129 GluLysAlaAspHisGluGluLeuCysGluPheArgProTyrSerCysProCysProGly 148
      |||||
Db      523 GAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCTTTATTCCTGTCGTCGCCCTGGT 582
Qy      149 AlaSerCysLysTrpGlnGlySerLeuAspAlaValMetProHisLeuMetHisGlnHis 168
      |||||
Db      583 GCTTCTCTGTAATGGCAGGCTCTCTGGATGCTGTATGCCCCATCTGATGCATCAGCAT 642
Qy      169 LysSerIleThrThrLeuGlnGlyGluAspIleValPheLeuAlaThrAspIleAsnLeu 188
      |||||
Db      643 AAGTCCATTACAACCTACAGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTT 702
Qy      189 ProGlyAlaValAspTrpValMetMetGlnSerCysPheGlyPheHisPheMetLeuVal 208
      |||||
Db      703 CCTGGTGTCTGTGACTGGGTGATGATGCAGTCTCTGTTTGGCTTTTCACTTCATGTGTAGTC 762
Qy      209 LeuGluLysGlnGluLysTyrAspGlyHisGlnPhePheAlaIleValGlnLeuIle 228
      |||||
Db      763 TTAGAGAAACAGAAATAATACATGTTGCCAGAGTCTTTCGAATCGTACAGCTGATA 822
Qy      229 GlyThrArgLysGlnAlaGluAsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArg 248
      |||||
Db      823 GGAACACGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGCTCATAGCGCA 882
Qy      249 ArgLeuThrTrpGluAlaThrProArgSerIleHisGluGlyIleAlaThrAlaIleMet 268
      |||||
Db      883 CGATTGACTTGGGAAGCACTCTCGATCTATTTCATGAGGAATTCACAGCACTTATG 942
Qy      269 AsnSerAspCysLeuValPheAspThrSerIleAlaGlnLeuPheAlaGluAsnGlyAsn 288
      |||||
Db      943 AATAGCGACTGTCTAGTCTTTGACACAGCATTCACAGCTTTTTCGAGAAAATGGCAAT 1002
Qy      289 LeuGlyIleAsnValThrIleSerMetCys 298
      |||||
Db      1003 TTAGGATCAATGTAATCTATTTCATGTGT 1032
      |||||

RESULT 14
ID ADS34359
AC ADS34359 standard; DNA; 2454 BP.
XX
AC ADS34359;
XX
DT 02-DEC-2004 (first entry)
DE
DE POSH protein associated DNA #113.
XX
KW ds; gene; cytostatic; nontropic; neuroprotective; antiparkinsonian;
KW anticonvulsant; antiviral; neuroleptic; central nervous system;
KW POSH polypeptide; POSH-associated protein; POSH-AP; HERPUDI;
KW Ubiquitin ligase; antiviral agent; anti-apoptotic agent;
KW anti-cancer agent; secretory pathway trafficking inhibitor;
KW neurological disorder progression disorder; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; schizophrenia;
KW Niemann-Pick's disease.
XX

```

```

OS Homo sapiens.
XX
PN WO2004078130-A2.
XX
PD 16-SEP-2004.
XX
PF 02-MAR-2004; 2004WO-US006308.
XX
PR 03-MAR-2003; 2003US-0451437P.
PR 05-MAR-2003; 2003US-0452284P.
PR 19-MAR-2003; 2003US-0455760P.
PR 20-MAR-2003; 2003US-0456640P.
PR 03-APR-2003; 2003US-0460526P.
PR 04-APR-2003; 2003US-0460792P.
PR 21-APR-2003; 2003US-0464285P.
PR 09-MAY-2003; 2003US-0469462P.
PR 15-MAY-2003; 2003US-0471378P.
PR 20-MAY-2003; 2003US-0472327P.
PR 30-MAY-2003; 2003US-0474706P.
PR 03-JUN-2003; 2003US-0475825P.
PR 17-JUN-2003; 2003US-0479317P.
PR 19-JUN-2003; 2003US-0480215P.
PR 19-JUN-2003; 2003US-0480376P.
PR 08-AUG-2003; 2003US-0493860P.
PR 28-AUG-2003; 2003US-0498634P.
PR 16-SEP-2003; 2003US-0503931P.
PR 10-NOV-2003; 2003WO-US035712.
PR 05-FEB-2004; 2004WO-US003600.
PR 02-MAR-2004; 2004US-0549896P.
XX
PA (PROT-) PROTEOLOGICS INC.
XX
PI Taglicht DN, Alroy I, Reiss Y, Yaar L, Ben-Avraham D, Tuvia S;
PI Greener T;
XX
XX WPI; 2004-652346/64.
XX
PT Isolated, purified or recombinant complex, useful for identifying an
PT antiviral, anti-apoptotic or anti-cancer, comprising POSH polypeptide and
PT POSH-associated protein (POSH-AP).
XX
PS Disclosure; SEQ ID NO 123; 374pp; English.
XX
CC The invention relates to an isolated, purified or recombinant complex (I)
CC comprising a POSH polypeptide and a POSH-associated protein (POSH-AP) (a)
CC or HERPUDI and a Ubiquitin ligase (b). Methods using (I), (a) or (b) are
CC useful for identifying an agent that modulates an activity of a POSH
CC polypeptide or POSH-AP, for identifying an antiviral agent, an anti-
CC apoptotic agent, an anti-cancer agent, an agent that inhibits trafficking
CC of a protein through the secretory pathway, an agent that inhibits the
CC progression of a neurological disorder, an agent that modulates a POSH
CC function, an agent that modulates a HERPUDI function. The methods can be
CC used for treating a viral infection, for inhibiting an activity of a POSH
CC -AP in a cell, for treating a POSH-associated disease in a subject. The
CC POSH-associated disease is viral infection, POSH-associated cancer or
CC POSH-associated neurological disorder. The methods are useful for
CC treating or preventing POSH-associated neurological disorder in a subject
CC e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease,
CC schizophrenia, Niemann-Pick's disease. This sequence corresponds to a
CC nucleic acid of the invention.
XX
SQ Sequence 2454 BP; 743 A; 456 C; 510 G; 745 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 1-69e-131 Length: 2454
Score: 1537.00 Matches: 283
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.41% Indels: 0
DB: 13 Gaps: 0

```

US-10-679-246-2 (1-298) x ADS34359 (1-2454)

QY	16	GluMetSerArgGlnThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSer	35
DB	457	GAATGAGCGCTCAGACTGTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACCATCC	516
QY	36	GlnArgValProAlaLeuThrGlyThrAlaSerAsnAsnAspLeuAlaSerLeuPhe	55
DB	517	CAGAGGGTGGCTCCCTGACTGGCACCACTGCATCCACCAATGACTTGGCGAGTCTTTT	576
QY	56	GluCysProValCysPheAspTyrValLeuProProLeuLeuGlnCysGlnSerGlyHis	75
DB	577	GAGTGTCCAGTCTGCTTTGACTATGTGTGTACCGCCATTTCTTCAATGTGAGAGTGGCCAT	636
QY	76	LeuValCysSerAsnCysArgProLysLeuThrCysCysProThrCysArgGlyProLeu	95
DB	637	CTTGTTGTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCAATGCGGGGCCCTTTG	696
QY	96	GlySerIleArgAsnLeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLys	115
DB	697	GGATCCATTGCAACTTGGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAA	756
QY	116	TyrAlaSerSerGlyCysGluLeuThrLeuProHisThrGluLysAlaAspHisGluGlu	135
DB	757	TATGCGTCTTCTGGATGTGAATAAATCTGCCACACACAGAAAAGCAGACCATGAAGAG	816
QY	136	LeuCysGluPheArgProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGly	155
DB	817	CTCTGTGAGTTTAGGCTTATTCCTGTCCGTGCGCTGTGCTTCTGTAAATGGCAAGC	876
QY	156	SerLeuAspAlaValMetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGln	175
DB	877	TCTCTGATGCTGTATGATGCCCATCTGATGCATCAGCATAAGTCCATTACAACCCCTACAG	936
QY	176	GlyGluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrVal	195
DB	937	GGAGAGGATATAGTTTCTTGTGTACAGACATTAATCTTCCGTGCTGTGACTGGGTG	996
QY	196	MetMetGlnSerCysPheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyr	215
DB	997	ATGATGCGTCTCTTTTGGCTTTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT	1056
QY	216	AspGlyHisGlnGlnPheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGlu	235
DB	1057	GATGGTCACAGCAGTCTTTCGCAATCGTACAGCTGATAGGAACAGCAAGCAAGCTGAA	1116
QY	236	AsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArgLeuThrTyrGluAlaThr	255
DB	1117	AATTTTGTCTTACCGACTTGAGCTTAATGTGTATAGGCGAGCATTTGCTGGAGCGACT	1176
QY	256	ProArgSerIleHisGluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPhe	275
DB	1177	CCTCGATCTATTATGAGGAATTCGAACAGCCATTATGAATAGCGACTGTCTAGTCTTT	1236
QY	276	AspThrSerIleAlaGlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIle	295
DB	1237	GACACACGACTTGACAGCTTTTTCGAGAAAATGGCAATTTAGGCATCAATGAATATT	1296
QY	296	SerMetCys	298
DB	1297	TCCATGTGT	1305
RESULT	15		
ID	ABI99429		
XX	ABI99429 standard; cDNA; 1968 BP.		
XX	ABI99429;		
DT	07-MAR-2002 (first entry)		
XX	Mouse ischaemic condition related cdna sequence SEQ ID NO:354.		
DE	Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;		
XX	vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.		
XX			
OS	Mus musculus.		
XX	W0200188188-A2.		
PN	22-NOV-2001.		
PD	18-MAY-2001; 2001WO-JP004192.		
PF	18-MAY-2000; 2000JP-00145977.		
XX	(UNNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.		
PA	Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;		
PI	WPI; 2002-034733/04.		
XX	P-P8DB; ABB57146.		
DR	Examining the ischemic condition (e.g. occlusive ischemia) by measuring		
XX	expression levels of particular genes defined in the specification or by		
PT	determining the expression profile of a gene group comprising these		
PT	genes.		
PT	Claim 2; Page 956-958; 2690pp; English.		
XX	The present invention describes a method for examining ischaemic		
CC	conditions, comprising measuring the expression levels of particular		
CC	genes (I) in a test sample or determining the expression profile of a		
CC	gene group in the sample comprising genes selected from (I). The method		
CC	is useful for examining the ischaemic condition (e.g. compressive		
CC	ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring		
CC	expression levels of particular genes (ABI99202 to ABI99912, encoding the		
CC	protein sequences in ABB57020 to ABB57374) or by determining the		
CC	expression profile of a gene group comprising these genes. The expression		
CC	levels or expression profiles produced by these genes are used as an		
CC	indicator when screening for ischaemic condition-improving drugs or		
CC	therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR		
CC	primers for a mouse ischaemic condition related sequence, which are used		
CC	in the exemplification of the present invention		
XX	Sequence 1968 BP; 500 A; 397 C; 447 G; 624 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	1.55e-131	Length:	1968
Score:	1536.00	Matches:	284
Percent Similarity:	98.62%	Conservative:	2
Best Local Similarity:	97.93%	Mismatches:	0
Query Match:	95.34%	Indels:	4
DB:	6	Gaps:	1
US-10-679-246-2 (1-298) x ABI99429 (1-1968)			
QY	9	ProProTyrValPheIleSerGluMetSerArgGlnThrAlaThrAlaLeuProThrGly	28
DB	109	CGCCCC-----ACAGAGATGAGCGCCAGACTGTCTACAGCATTTACCCATGGC	156
QY	29	ThrSerLysCysProProSerGlnArgValProAlaLeuThrGlyThrThrAlaSerAsn	48
DB	157	ACCTCAAAGTGTCCACCATCCCGAGGGGTACTGCTTGACCGGCACAACTGATCCAAC	216
QY	49	AsnAspLeuAlaSerLeuPheGluCysProValCysPheAspTyrValLeuProPhe	68
DB	217	AATGACTTGGCGAGCTTTTGTGAGTGTCTGCTGCTTGTGCTTGTGCTTGTGCTTGT	276
QY	69	LeuGlnCysGlnSerGlyHisLeuValCysSerAsnCysArgProLysLeuThrCysCys	88
DB	277	CTTCAGTGTGAGAGTGGCCATCTTGTGTTGTAGCACTGTGCGCCCAACTTACATTTGT	336
QY	89	ProThrCysArgGlyProLeuGlySerIleArgAsnLeuAlaMetGluLysValAlaAsn	108
DB	337	CCCACTTGGCGGGGCCCATTTGGGATCCATTCGCACTTGGCTATGGAGAAAGTGGCAAC	396
QY	109	SerValLeuPheProCysLysTyrAlaSerSerGlyCysGluLeuThrLeuProHisThr	128

397	TCAGTACTCTCCCTTGTAAATATATGCTCTCTGGATGTGAATAAATACTCTGCCACACACC	456
129	GlulysAlaAspHisGluGluLeuCysGluPheArgProTyrSerCysProCysProGly	148
457	GAAGAAGGCAGACACGAGGAGCTCTGTGAGTTTCAGGSCCTTACTCTGCCCCCTGCCTGGT	516
149	AlaSerCysLysTrpGlnGlySerLeuAspAlaValMetProHisLeuMethHisGlnHis	168
517	GCCTTCCTGTAAGTGCGCAAGGCTCCTTGGATGCGGTATGCCCCACCTGATGCATCAGCAC	576
169	LysSerIleThrThrLeuGlnGlyGluAspIleValPheLeuAlaThrAspIleAsnLeu	188
577	AAGTCCATTACACCCTCAAGGAGGAATATAGTATTTCTCTTGCTACAGACATTAAACCTT	636
189	ProGlyAlaValAspTrpValMetMetGlnSerCysPheGlyPheHisPheMetLeuVal	208
637	CCTGGTGTGTTGACTGGTGATGATGCAGCTTGCTTTTGGCTTTCAATTCATGTTAGTC	696
209	LeuGluLysGlnGluLysTyrAspGlyHisGlnGlnPhePheAlaIleValGlnLeuIle	228
697	TTGAGAGAAACAAGAAAAATATGATGGTCATCAGCAGTTCCTTTTGCAATTTGTACAAC	756
229	GlyThrArgLysGlnAlaGluAsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArg	248
757	GGACACGCAAGCAAGCTGGAATTTTTCGATATCGACTTGAGCTAAATTTGGTCATAGGCGG	816
249	ArgLeuThrTrpGluAlaThrProArgSerIleHisGluGlyIleAlaThrAlaIleMet	268
817	CGATTGACTTGGGAAGCGCACTCCTCGGTCTATTCAATGAGGGAAATTTGCAACGCCCA	876
269	AsnSerAspCysLeuValPheAspThrSerIleAlaGlnLeuPheAlaGluAsnGlyAsn	288
877	AATAGTGACTGCCTAGTGTGTTTGACACCAAGCATTTGCACAGCTTTTTCGAGAAAT	936
289	LeuGlyIleAsnValThrIleSerMetCys	298
937	TTAGGCATCAATGTAACTATTTCATGTGT	966

Search completed: April 25, 2005, 07:02:25  
Job time : 639 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 25, 2005, 06:43:10 ; Search time 3685 Seconds  
(without alignments)  
3078.196 Million cell updates/sec

Title: US-10-679-246-2

Perfect score: 1611

Sequence: 1 MWIIIFLLPPYVFISEMSRQ.....IAQLFAENGLGINTWISC 298

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US10679246/runat.22042005.185510.7697/app.query.fasta\_1.455  
-DB=EST -OPWT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCAUGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10679246 @CNC 1.1 3437 @runat.22042005.185510.7697 -NCPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hic:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gsa1:  
9: gb\_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	1221	7 U70056	U70056 U70056 Soar
2	1537	95.4	1360	3 CR607323	CR607323 full-length
3	1537	95.4	1857	3 CR614925	CR614925 full-length
4	1532	95.1	849	9 AY402255	AY402255 Homo sapi
5	1528	94.8	849	9 AY402257	AY402257 Mus muscu
6	1521	94.4	849	9 AY402256	AY402256 Pan trogl
7	1465.5	91.0	859	5 BX366404	BX366404
8	1427	88.6	1287	3 CR598465	CR598465 full-length
9	1412	87.6	1180	7 CK025134	CK025134 AGENCOURT

10	1360	84.4	848	5	BU614872	BU614872
11	1325	82.2	824	4	BG619351	BG619351
12	1319	81.9	921	5	BX355841	BX355841
13	1285	79.8	872	5	BX448317	BX448317
14	1269	78.8	716	5	BU260826	BU260826
15	1247	77.4	779	1	AI936347	AI936347
16	1247	77.4	852	5	BX463976	BX463976
17	1229	76.3	890	7	CF551876	CF551876
18	1209	75.0	806	7	CF752915	CF752915
19	1207	74.9	849	5	BU912024	BU912024
20	1206	74.9	2243	3	CR610024	CR610024
21	1196	74.2	882	5	BX757871	BX757871
22	1195	74.2	785	5	BX384656	BX384656
23	1180	73.9	990	6	BY761918	BY761918
24	1180	73.2	789	7	CV116728	CV116728
25	1175	72.9	666	1	AL656230	AL656230
26	1170	72.6	953	6	CA980024	CA980024
27	1168	72.5	649	2	BF446537	BF446537
28	1162	72.1	778	1	AI971258	AI971258
29	1161	72.1	884	7	CR584650	CR584650
30	1154	71.6	717	5	BU261233	BU261233
31	1146	71.1	622	7	CV030563	CV030563
32	1146	71.1	800	4	BI836179	BI836179
33	1129	70.1	621	7	CN677231	CN677231
34	1129	70.1	647	2	BES03207	BES03207
35	1119	69.5	647	2	BES02849	BES02849
36	1112	69.0	903	5	BX770276	BX770276
37	1111.5	69.0	844	7	CO804390	CO804390
38	1111.5	69.0	885	5	BU164429	BU164429
39	1103	68.5	734	1	AU138740	AU138740
40	1099.5	68.2	842	5	BX459114	BX459114
41	1099	68.2	811	2	AW128457	AW128457
42	1093	67.8	902	4	BI763325	BI763325
43	1086	67.4	735	4	BI560892	BI560892
44	1084.5	67.3	715	5	BX459079	BX459079
45	1084.5	67.3	769	5	BX097203	BX097203

ALIGNMENTS

U70056 1221 bp mRNA linear EST 27-OCT-1999  
U70056 Soares infant brain INIB Homo sapiens cDNA clone 25050, mRNA sequence.

U70056  
U70056.1 GI:2731404  
EST.  
EST.  
Homo sapiens (human)  
Homo sapiens

REFERENCE  
1 (bases 1 to 1221)  
Volorio, S., Simon, G., Repetto, M., Cucciard, M., Banfi, S., Boran, G., Ballabio, A. and Zollo, M.  
Sequencing analysis of forty-eight human image cDNA clones similar to drosohila mutant protein  
DNA Seq. 9 (5-6), 307-315 (1998)

JOURNAL  
MEDLINE  
PUBMED  
99452388  
10524757

COMMENT  
Contact: Zollo, Massimo  
Telethon Institute of Genetics and Medicine  
Via Olgettina 58, Milan, MI 20132, Italy  
Email: zollo@tigem.it.

FEATURES  
source  
1..1221  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="16q12.1-q13"  
/clone="25050"  
/sex="female"  
/dev\_stage="73 days post natal"



Db 183 CTGTTTGTAGTACCTGTCCTCCAAAGCTCACAATGTGTCACATCTGCGGGCCCTTTG 242  
 Qy 96 GlySerIleArgAsnLeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLys 115  
 Db 243 GGATCCATTCGCACTTGGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAA 302  
 Qy 116 TyrAlaSerSerGlyCysGluIleThrLeuProHisThrGluLysAlaAspHisGluGlu 135  
 Db 303 TATGCGTCTTCTGGATGTGAATAAATCTGCCACACACAGAAAAGCAGACCATGAAGAG 362  
 Qy 136 LeuCysGluPheArgProTyrSerCysProCysProGlyValaSerCysLysTrpGlnGly 155  
 Db 363 CTCTGTAGTTAGGCTTATCTCTGCTGCGCCCTGGTCTTCTGTAATGGCAAGGC 422  
 Qy 156 SerLeuAspAlaValMetProHisLeuMetHisGlnHisLysSerIleThrLeuGln 175  
 Db 423 TCTCTGGATGCTGTAATGCCGATCTGATGCATCAGCAATAAGTCCATTACACCCCTACAG 482  
 Qy 176 GlyGluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAlaValaAspTrpVal 195  
 Db 483 GGAGAGGATATAGTTTCTTGTCTACAGACATTAATCTTCTGCTGCTGTGACTGGGTG 542  
 Qy 196 MetMetGlnSerCysPheGlyPheHisLeuMetLeuValLeuGluLysGlnGlySerTyr 215  
 Db 543 ATGATCGAGTCTGTTTGGCTTTTCACTTCATGTTAGTCTTATAGAGAAACAGGAAAAATAC 602  
 Qy 216 AspGlyHisGlnGlnPhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGlu 235  
 Db 603 GATGTCACACAGAGTCTTCTGCAATCTGATAGTATAGGAAACAGCAGCAGCTGAA 662  
 Qy 236 AsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArgArgLeuThrTrpGluAlaThr 255  
 Db 663 AATTTTGTCTACCGACTTAAATGCTATAGTCTGATAGGCAAGTCTGCTAGTCTTT 722  
 Qy 256 ProArgSerIleHisGluGlyLeuAlaThrAlaIleMetAsnSerAspCysLeuValPhe 275  
 Db 723 CTCTCGATCTATTATCATGAAGAAATGCAACAGCCATTATGAATAGCCACTGTCTAGTCTTT 782  
 Qy 276 AspThrSerIleAlaGlnLeuPheAlaGluAsnGlyAsnValThrIle 295  
 Db 783 GACACAGCATGTCACAGCTTTTGGAGAAATGGCAATTTAGGCATCAATGTAATATT 842  
 Qy 296 SerMetCys 298  
 Db 843 TCCATGTGT 851

## RESULT 3

## CR614925

## LOCUS

## DEFINITION

1857 bp mRNA linear HTC 21-JUL-2004

full-length cDNA clone CS0DK012Y120 of HeLa cells Cot 25-normalized

of Homo sapiens (human).

CR614925.1 GI:50495732

HTC; CDSLT CDNA.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Palayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
2 (bases 1 to 1857)

REFERENCE

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

## FEATURES

## source

1..1857  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DK012Y120"  
 /tissue\_type="HeLa cells Cot 25-normalized"  
 /plasmid="pCMVSPORT\_6"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,06e-137 Length: 1857  
 Score: 1537.00 Matches: 283  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 95.41% Indels: 0  
 DB: 3 Gaps: 0

US-10-679-246-2 (1-298) x CR614925 (1-1857)

Qy 16 GluMetSerArgGlnThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSer 35  
 Db 3 GAAATGAGCGCTGACACTGCTACAGCATTTACCTACCGGTACCTCGAAGTGTCCACCATCC 62  
 Qy 36 GlnArgValProAlaLeuThrGlyThrAlaSerAsnAsnAspLeuAlaSerLeuPhe 55  
 Db 63 CAGAGGTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 122  
 Qy 56 GluCysProValCysPheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHis 75  
 Db 123 GAGTGTCCAGTCTGCTTTCATGTTATGTTACGCCCATCTTCAATGTCCAGATGGCCAT 182  
 Qy 76 LeuValCysSerAsnCysArgProLysLeuThrCysCysProThrCysArgGlyProLeu 95  
 Db 183 CTTGTTTGTAGCAACTGTGCGCAAGCTCACAATGTTGTCCAACCTGCGGGGCCCTTTG 242  
 Qy 96 GlySerIleArgAsnLeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLys 115  
 Db 243 GGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTTAATTCAGTACTTTTCCCTGTAAA 302  
 Qy 116 TyrAlaSerSerGlyCysGluIleThrLeuProHisThrGluLysAlaAspHisGluGlu 135  
 Db 303 TATGCGTCTTCTGGATGTGAATAAATCTGTCACACACAGAAAAGCAGCAGCATGAAGAG 362  
 Qy 136 LeuCysGluPheArgProTyrSerCysProCysProGlyAlaSerCysLysTrpGlnGly 155  
 Db 363 CTCTGTAGTCTTAGGCTTATCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422  
 Qy 156 SerLeuAspAlaValMetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGln 175  
 Db 423 TCTCTGGATGCTGTAATGCCCATCTGATGCATCAGCAATAAGTCCATTACACCCCTACAG 482  
 Qy 176 GlyGluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAlaValaAspTrpVal 195  
 Db 483 GGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCT 542  
 Qy 196 MetMetGlnSerCysPheGlyPheHisPheMetLeuValLeuGluLysGlnGlySerTyr 215  
 Db 543 ATGATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602  
 Qy 216 AspGlyHisGlnGlnPhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGlu 235  
 Db 603 GATGTCACACAGAGTCTTCCGAATTCGATCAGTATAGGAAACAGCAGCAGCATGAAGAG 662  
 Qy 236 AsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArgArgLeuThrTrpGluAlaThr 255  
 Db 663 AATTTTGTCTACCGACTTAAATGCTATAGTCTGATAGGCAAGTCTGCTAGTCTTT 722  
 Qy 256 ProArgSerIleHisGluGlyLeuAlaThrAlaIleMetAsnSerAspCysLeuValPhe 275





Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.

## TITLE

Direct Submission  
JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA

COMMENT  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment

## FEATURES

source  
1. 849  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..849  
/gene="Slah1"  
/locus\_tag="HCM1163"

## ORIGIN

Alignment Scores:  
Pred. No.: 6,956-137 Length: 849  
Score: 1528.00 Matches: 281  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.65% Mismatches: 0  
Query Match: 94.85% Indels: 0  
DB: 9 Gaps: 0

US-10-679-246-2 (1-298) x AY402257 (1-849)

Qy 17 MetSerArgGlnThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGln 36  
Db 1 ATGAGCGCGCAGACTGCTACAGATTACCACCTGGCAGCTCAAGTGTCACCATCCAG 60  
Qy 37 ArgValProAlaLeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGlu 56  
Db 61 AGGCTACCTGCTGACCGGCACTGCAATGCATCAACATGACTGGCAGTCTTTTGAG 120  
Qy 57 CysProValCysPheAspTyrValLeuProPheLeuGlnCysGlnSerGlyHisLeu 76  
Db 121 TGTCTCTGCTGCTTGTGACTATGTTGCCACCTATTCTTCAGTGTGAGAGTGCACATCT 180  
Qy 77 ValCysSerAsnCysArgProLysLeuThrCysProThrCysArgGlyProLeuGly 96  
Db 181 GTTGTAGCACTGTCGCCCAACTTACATGTTGCCACCTGCGGGGCCCATTTGGA 240  
Qy 97 SerLeuArgAsnLeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyr 116  
Db 241 TCCATTGCAACTGCTGATGAGAAAGTGGCAACTCAGTACTCTTCCTTTGTAATAT 300  
Qy 117 AlaSerSerGlyCysGluLeuThrLeuProHisThrGluLysAlaAspHisGluLeu 136  
Db 301 GCCTCTCTGGAATGTAATAACTCTGCCACACACCGAAAGGCAGACGAGGAGCTC 360  
Qy 137 CysGluPheArgProTyrSerCysProCysProGlyValaSerCysLysTyrGlnGlySer 156  
Db 361 TGTGAGTTCAGGCTTACTCTGCGCCCTGCTGCTGCTTCTGTAAGTGGCAAGCTCC 420  
Qy 157 LeuAspAlaValMetProHisLeuMetHisGlnHisLysSerLeuThrThrLeuGlnGly 176  
Db 421 TTGGATGCGCTCATGCCACCTGATGATCAGCAAGTCAATACCACTCCGCAAGA 480  
Qy 177 GluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyValaValAspTyrValMet 196  
Db 481 GAAGATATATAGTTTCTGCTACAGACATTAACCTTCCTGGTGTGTTGACTGGGTGATG 540  
Qy 197 MetGlnSerCysPheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAsp 216  
Db 541 ATGAGCTCTGTTTGGCTTTTCAATTTTCATGTTAGTCTTGGAGAAACAGAAATATGAT 600  
Qy 217 GlyHisGlnGlnPhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsn 236  
Db 601 GGTGATCAGCAGTCTTTTGCATTTGTACAACTGTACAGTAAGAACACGCAAGCACTGAAAT 660  
Qy 237 PheAlaTyrArgLeuGlnLeuAsnGlyHisArgArgLeuThrTyrTrpGluAlaThrPro 256

Db 661 TTGTCATATCGACTTGGCTAAATGGTCATAGCGCGGATTGACTTGGGAGCGACTCCT 720  
Qy 257 ArgSerIleHisGluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAsp 276  
Db 721 CGGTCTATTTCATGAGGGAATTCACAGCAATTTATGAATAGTACTGCTAGTGTGAC 780  
Qy 277 ThrSerIleAlaGlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSer 296  
Db 781 ACCAGCATTCACAGCTTTTTCGAGAAATGGCAATTTAGGCATCAATTAATATTTCC 840  
Qy 297 MetCys 298  
Db 841 ATGTGT 846

RESULT 6  
AY402256 849 bp DNA linear GSS 15-DEC-2003  
LOCUS  
DEFINITION  
Pan troglodytes Slah1 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION  
AY402256  
VERSION  
AY402256.1 GI:39758242  
KEYWORDS  
GSS.  
SOURCE  
Pan troglodytes (chimpanzee)  
ORGANISM  
Pan troglodytes

REFERENCE  
AUTHORS  
Clark, A.G., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.

TITLE  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL  
Science 302 (5652), 1960-1963 (2003)

REFERENCE  
AUTHORS  
Clark, A.G., Glanowski, S., Nielsen, R., Thomas, P., Kejariwal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.

TITLE  
JOURNAL  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.

## FEATURES

source  
1. 849  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..849  
/gene="Slah1"  
/locus\_tag="HCM1163"

## gene

## ORIGIN

Alignment Scores:  
Pred. No.: 3,3e-136 Length: 849  
Score: 1521.00 Matches: 281  
Percent Similarity: 99.65% Conservative: 0  
Best Local Similarity: 99.65% Mismatches: 1  
Query Match: 94.41% Indels: 0  
DB: 9 Gaps: 0

US-10-679-246-2 (1-298) x AY402256 (1-849)

Qy 17 MetSerArgGlnThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGln 36  
Db 1 ATGAGCGCGCAGACTGCTACAGATTACCACCTGGCAGCTCAAGTGTCACCATCCAG 60  
Qy 37 ArgValProAlaLeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGlu 56  
Db 61 AGGCTACCTGCTGACCGGCACTGCAATGCATCAACATGACTGGCAGTCTTTTGAG 120





Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Wen Zon, Harvard  
CDNA Library Preparation: Open Biosystems  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: ZGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM14819 row: h column: 07  
High quality sequence stop: 877.  
Location/Qualifiers  
1. .1180  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/db\_xref="taxon:7955"  
/clone="IMAGE:7049721"  
/tissue\_type="whole body"  
/lab\_host="DH10B"  
/clone\_lib="NH ZGC 10"  
/notes="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;  
Bulk tissue was collected from a whole adult individual  
from the Tuebingen strain. 1st strand cDNA was primed with  
a Not I - oligo(dT) primer, double-stranded cDNA was  
cloned into the Not I and EcoRV sites of pExpress-1.  
Library was size-selected for >1 kb fragments. A  
normalized version of this library is also available  
(NH ZGC 7). Library was constructed by Open Biosystems  
(Huntsville, AL)."

## FEATURES source

```

/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7049721"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV; Bulk tissue was collected from a whole adult individual from the Tuebingen strain. 1st strand cDNA was primed with a Not I - oligo(dT) primer, double-stranded cDNA was cloned into the Not I and EcoRV sites of pExpress-1. Library was size-selected for >1 kb fragments. A normalized version of this library is also available (NIH_ZGC_7). Library was constructed by Open Biosystems (Huntsville, AL)."
```

ORIGIN

Alignment Scores:	
Pred. No.:	1.82e-125
Score:	1412.00
Percent Similarity:	99.62%
Best Local Similarity:	97.35%
Query Match:	87.65%
DB:	7

Length:	1180
Matches:	257
Conservative:	6
Mismatches:	1
Indels:	0
Gaps:	0

HS-10-679-246-2 (1-298) x CK025134 (1-1180)

## ORIGIN

Alignment Scores:			
Pred. No.:	1..92e-125	Length:	1180
Score:	1432.00	Matches:	257
Percent Similarity:	99.62%	Conservative:	6
Best Local Similarity:	97.35%	Mismatches:	1
Query Match:	87.65%	Indels:	0
DB:	7	Gaps:	0
US-10-679-246-2 (1-298) x CK025134 (1-1180)			
Qy	35	SerGlnArgValProAlaLeuThrGlyThrAlaSerAsnAspLeuAlaSerLeu	54
Db	3	TCCTCAGCGGTTCCAACTTTGTGGGACACACACATCCCAACAGTGACCTGCCAGCGCTG	62
Qy	55	PheGluCysProValCysPheAspTyrValLeuProProIleLeuGlnCysGlnSerGly	74
Db	63	TTTGAGTGCCGGGTCTGTTTCGACTATGTGTGCGGCCCATCTCTGCAGTCCAGAGCGGC	122
Qy	75	HisLeuValCysSerAsnCysArgProIlysLeuThrCysCysProThrCysArgGlyPro	94
Db	123	CATTGTGTGTAGCAACTCGCGGCTAAATCACTCTGTCGCCACCTGCAGAGGCGCG	182
Qy	95	LeuGlySerIleArgAsnLeuAlaMetGluIysValAlaAsnSerValLeuPheProCys	114
Db	183	CTTGGGTGCATCCGTAAATTTGGCAATGGAGAAAGTGCCCACTCAGTGTCTTTCCCTGCG	242
Qy	115	LysTyrAlaSerSerGlyCysGluIleThrLeuProHisThrGluIysAlaAspHisGlu	134
Db	243	AAGTACGCTCATCGGGCTGCGAAGTCATCTACCGCACACACAAAGCTGAACACGAG	302
Qy	135	GluLeuCysGluPheArgProTyrSerCysProCysProGlyAlaSerCysLeuThrGln	154
Db	303	GAGCTTTGTGAATTTCCGGCCATATCTTGTGCCCTGTCTCTGGCGGCTCTCGAAGTGGAG	362
Qy	155	GlySerLeuAspAlaValMetProHisLeuMetHisGlnHisSerIleThrThrLeu	174
Db	363	GGATCTCTCGAGCGGTTATGCCCACTTGTGTGCCACGACCAAGTCCATAACCACTA	422
Qy	175	GlnGlyGluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrp	194
Db	423	CAAGGTGAGGATATAGTCTTCTGTGCCACACACATCAACTTGTCCCTGGAGCGAGTGGACTGG	482
Qy	195	ValMetMetGlnSerCysPheGlyPheHisPheMetLeuValLeuGluIysGlnGluIys	214

Db	483	GT	CAT	GAT	G	CA	A	T	C	T	G	C	T	T	C	G	C	T	T	C	A	T	T	T	A	T	G	T	G	T	G	G	A	A	G	A	A	542	
Qy	215	Tyr	Asp	Gly	His	Gln	Gln	Phe	Ala	Ile	Val	Gln	Leu	Ile	Gly	Thr	Arg	Lys	Gln	Ala	234																		
Db	543	TAG	CAG	CGC	CAC	CA	CAG	AT	T	C	T	T	T	T	G	C	A	C	T	C	A	A	T	G	C	G	C	A	C	G	A	A	C	A	A	G	C	602	
Qy	235	Glu	Asn	Phe	Ala	Tyr	Arg	Leu	Glu	Leu	Asn	Gly	His	Arg	Arg	Leu	Thr	Trp	Glu	Ala	254																		
Db	603	GAA	AA	C	T	T	C	C	T	T	A	A	C	G	G	C	C	A	G	G	C	G	G	C	A	C	T	C	A	C	T	T	G	G	A	G	C	662	
Qy	255	Thr	Pro	Arg	Ser	Ile	His	Glu	Gly	Ile	Ala	Thr	Ala	Ile	Met	Ile	Ser	Asp	Cys	Leu	Val	274																	
Db	663	ACG	CA	C	G	T	T	C	T	A	T	C	A	T	G	A	G	G	C	A	T	C	A	T	G	A	C	A	T	G	A	C	A	T	G	T	G	T	722
Qy	275	Phe	Asp	Thr	Ser	Ile	Ala	Gln	Leu	Phe	Ala	Glu	Asn	Gly	Asn	Leu	Gly	Ile	Asn	Val	Thr	294																	
Db	723	TTT	GAC	ACC	T	C	T	A	T	T	G	T	C	A	G	T	T	T	G	C	G	A	A	C	G	G	C	A	C	T	T	A	C	T	C	A	C	782	
Qy	295	Ile	Ser	Met	Cys	298																																	
Db	783	ATA	T	C	A	T	A	T	G	T	G	C	794																										

RESULT 10

BU614872

LOCUS

DEFINITION

848 bp mRNA

linear EST 20-FEB-2003

UT-M-EVO-cbaf-j-08-0-UI.r1 NIH BMAP EV0 Mus musculus CDNA clone

UT-M-EVO-cbaf-j-08-0-UI 5', mRNA sequence.

## RESULT 10

BU614872	ACCESSION	REFERENCE
LOCUS	VERSION	AUTHORS
DEFINITION	KEYWORDS	TITLE
	SOURCE	JOURNAL
	ORGANISM	COMMENT

BU614872 848 bp mRNA linear EST 20-FEB-2003  
 UI-M-EV0-cbh-j-08-0-UI.r1 NIH BMAP EVO Mus musculus cDNA clone  
 UI-M-EV0-cbh-j-08-0-UI.5', mRNA sequence.

251. Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 848)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov  
Tissue procurement: Dr. James Lin, University of Iowa  
CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

(BMAP)  
The following repetitive elements were found in this cDNA  
sequence: 1-66, >GC-rich#Low\_complexity  
Seg primer: pyx-5.

FEATURES  
source

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-EVO-cbh-j-08-0-UI"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="OH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_EVO"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
Site 2: Not I; the library was constructed according to
Bonald, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with Ecor I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA

```

tail, is GTGGCTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

## ORIGIN

## Alignment Scores:

Pred. No.: 1-166-120 Length: 848  
Score: 1366.00 Matches: 254  
Percent Similarity: 96.97% Conservative: 2  
Best Local Similarity: 96.21% Mismatches: 4  
Query Match: 84.42% Indels: 5  
DB: 5 Gaps: 1

US-10-679-246-2 (1-298) x BU614872 (1-848)

QY 9 ProProTyrValPheLeuSerGluMetSerArgGlnThrAlaThrAlaLeuProThrGly 28  
DB 70 CCGCCC-----ACAGAGATGAGCGCGACAGCTGTACAGATTACCACTGGC 117  
QY 29 ThrSerLysCysProSerGlnArgValProAlaLeuThrGlyThrAlaSerAsn 48  
DB 118 ACCTCAAGTGTCCACCATCCAGAGGGTACCTGCTTGACCGCACACTGATCAAC 177  
QY 49 AsnAspLeuAlaSerLeuPheGluCysProValCysPheAspTyrValLeuProPhe 68  
DB 178 AATGACTTGGCGAGTCTTTTGAGTGTCTGTGCTTGTGACTATGTGTGGCACCTATT 237  
QY 69 LeuGlnCysGlnSerGlyHisLeuValCysSerArgCysArgProLysLeuThrCysCys 88  
DB 238 CTTCACTGTGAGAGTGGCCATCTTTGTGTAGCAACATGTCGCCCAACATTACATTTGT 297  
QY 89 ProThrCysArgGlyProLeuGlySerIleArgAsnLeuAlaMetGluLysValAlaAsn 108  
DB 298 CCCACTTGGCGGGCCCATTTGGATTCATTCGCACTTGGCTATGAGAAAGTGGCCAAC 357  
QY 109 SerValLeuPheProCysLysTyrAlaSerSerGlyCysGluIleThrLeuProHisThr 128  
DB 358 TCAGTACTCTTCCTTGTAAATATGCTCTCTTGGATGTGAATAACTCTGCCACACACC 417  
QY 129 GluLysAlaAspHisGluGluLeuCysGluPheArgProTyrSerCysProCysProGly 148  
DB 418 GAAAGCGAGACACAGAGAGCTCTGTGAGTTGAGCCCTTACTCTGCCCTGCCCTGGT 477  
QY 149 AlaSerCysLysTyrGlnGlySerLeuAspAlaValMetProHisLeuMetHisGlnHis 168  
DB 478 GCTTCTGTAAAGTGGCAAGGCTCTTGGATGCGCTCATGCCCCACCTGATCATCAAC 537  
QY 169 LysSerIleThrThrLeuGlnGlyGluAspIleValPheLeuAlaThrAspIleAsnLeu 188  
DB 538 AAGTCCATTACCACTGCAAGGAGAGATATAGTTTCTTCTGTACAGACATTAACTT 599  
QY 189 ProGlyAlaValAspTyrValMetMetGlnSerCysPheGlyPheHisPheMetLeuVal 208  
DB 598 CTTGGTGTCTGTGACTGGGTGATGATGATGCTTGTGTTGGCTTTCATTTGATGTC 657  
QY 209 LeuGluLysGlnGluLysTyrAspGlyHisGlnGlnPhePheAlaIleValGlnLeuIle 228  
DB 658 TTGGAGAAACAGANAATAATGATGTGTCATGATGATGATGATGATGATGATGATGAT 717  
QY 229 GlyThrArgLysGlnAlaGluAsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArg 248  
DB 718 GGNACAGCAAGCAAGCTGNAATTTTGCATATCGACTTGAGCTTAATGGTATATGCGG 777  
QY 249 ArgLeuThrTyrGluAlaThrProArgSerIleHisGluGlyIleAlaThrAlaIleMet 268  
DB 778 CGATTGACTTGGAGAGCACTCTCGTCTATTTCATGAGGGAATTGC-ACAGGCCCATATG 836  
QY 269 AsnSerAspCys 272  
DB 837 AATAGTGACTGC 848

## RESULT 11

BG619351

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG619351 824 bp mRNA linear EST 18-APR-2001  
602619484F1 NIH MGC 79 Homo sapiens cDNA clone IMAGE:4733228 5',  
mRNA sequence.  
BG619351  
BG619351.1 GI:13670722  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 824)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: CLONTECH Laboratories, Inc.  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Inyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLC1594 row: 9 column: 21  
High quality sequence stop: 748.

## FEATURES

source

1..824

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4733228"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NIH MGC 36"

/note="Organ: placenta; Vector: pDNR-LIB (Clontech);

Site 1: SfiI (ggccctggcc); Site 2: SfiI

(ggccatattggcc); 5' and 3' adaptors were used in cloning

as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'

and 3' adaptor sequence:

5'-ATTCTAGAGCGCGCGCCACATG-dT(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.3

kb (range 0.5-4.0 kb). 15/15 colonies contained inserts

by PCR. This library was enriched for full-length clones

and was constructed by Clontech Laboratories (Palo Alto,

CA). Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,666-117 Length: 824  
Score: 1325.00 Matches: 251  
Percent Similarity: 98.44% Conservative: 2  
Best Local Similarity: 97.67% Mismatches: 2  
Query Match: 82.25% Indels: 3  
DB: 4 Gaps: 0

US-10-679-246-2 (1-298) x BG619351 (1-824)

QY 43 GlyThrThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCysPheAsp 62  
DB 1 GGCACAACTGCATCCCAACATGACTTGGCGAGTCTTTTGGAGTGTCCAGTCTCTTGAC 60  
QY 63 TyrValLeuProPheLeuGlnCysGlnSerGlyHisLeuValCysSerAsnCysArg 82  
DB 61 TATGTGTACCGCCCATCTTCAATGTGAGTGGCCATCTTGTGTGACCACTGTCCG 120  
QY 83 ProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsnLeuAla 102  
DB 121 CCAAGCTCACATGTTGTCCAACTTGGCGGGCCCTTTGGGATCCATTCCAACTTGGCT 180  
QY 103 MetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGlyCysGlu 122  
DB 181 ATGGAGAAAGTGGCTAAATTCAGTACTTTTTCCCTGTAAATATGCGTCTTCTGGATGTGAA 240







```
RESULT 13
BX448317
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX448317 872 bp mRNA linear EST 06-MAY-2004
CS0DM009YG23 5-PRIME, mRNA sequence.
BX448317
BX448317.2 GI:47064237
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 872)
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequenced version replaced gi:31025931.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 8104.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cba?8=CS0AM009AD12QP1&c=8104.r.
Location/Qualifiers
1..872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DM009YG23"
/tissue type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Alignment Scores:
Pred. No.: 2,11e-113 Length: 872
Score: 1285.00 Matches: 244
Percent Similarity: 94.94% Conservative: 0
Best Local Similarity: 94.94% Mismatches: 12
Query Match: 79.76% Indels: 2
DB: 5 Gaps: 0

US-10-679-246-2 (1-298) x BX448317 (1-872)

QY 16 GluMetSerArgGlnThrAlaLeuProThrGlyThrSerLysCysProProSer 35
Db 3 GAAATGAGCGCTGCTACAGCTTACCTACCGTACCTCGAAGTGTCCACCATCC 62
QY 36 GlnArgValProAlaLeuThrGlyThrAlaSerAsnAsnAlaSerLeu-Ph 55
Db 63 CAGAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 122
QY 55 eGluCysProValCysPheAspTyrValLeuProProGlnCysGlnSerGlyHi 75
Db 123 TGAATATCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182
QY 75 sLeuValCysSerAsnCysArgProLysLeuThrCysCysProThrCysArgGlyProLe 95
Db 183 TCTTGTGTTAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
QY 95 uClySerIleArgAsnLeuAlaMetGlnLysValAlaAsnSerValLeuPheProCysLy 115
```

```
Db 243 GGGATCCATTGCGAAGTGGCTATGAGAAAGTGGCTAATTCATTCTTCCCTGTAA 302
QY 115 sTyrAlaSerSerGlyCysGluIleThrLeuProHisThrGluLysAlaAspHisGluGl 135
Db 303 ATATGCGTCTTCTGGATGTGAATAACTCTGCACACACAGAAAAAGCAGACCATGAAGA 362
QY 135 uLeuCysGluPheArgProTyrSerCysProCysProGlyAlaSerCysLysThrGlnGl 155
Db 363 GCTCTGTGAGTTAGGCTTATTCCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
QY 155 ySerLeuAspAlaValMetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGl 175
Db 423 CTCTCTGGATGCTGTATATGCCCTCATCTGATGATAGCATAGTCCATTACACCTTACA 482
QY 175 nGlyGluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpVa 195
Db 483 GGGAGAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCT 542
QY 195 lMetMetGlnSerCysPheGlyPheHisPheMetLeuValLeuGlnLysLysLysLys 215
Db 543 GATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
QY 215 rAspGlyHisGlnGlnPhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGl 235
Db 603 CGATGCTCACCAGCAGTCTTTCGCAATCGTACAGCTGATAGGACACGCAAGCAAGCTGA 682
QY 235 uAsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArgArgLeuThrTrpGluAlaTh 255
Db 663 AATTTTGTCTTACMGACTTGAGCBAATGCTCATAGCGCAGCATTTGACCTTGGGAAGCGMY 722
QY 255 rProArgSerIleHisGluGlyIleAlaThrAlaIleMetAsnSerAsp 271
Db 723 CCTCGGTGTATTCATGAAGGAATTGCAMA-GCCATTATGAATAGCGAY 770

RESULT 14
BU260826 716 bp mRNA linear EST 26-NOV-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

BU260826 603504644F1 CSEQCHN51 Gallus gallus cDNA clone CHEST427g12 5', mRNA
sequence.
BU260826
BU260826.1 GI:25528317
EST.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 716)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken CDNAs
Curr. Biol. 12 (22), 1965-1969. (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..716
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="CHEST427g12"
/dev_stage="22"
/lab_host="DH108"
/clone_lib="CSEQCHN51"
/note="Organ: limbs; Vector: pBluescript II KS(+); Site_1:
source
FEATURES
```



Db 252 TTTGACTATGTGTACCGCCCATCTTCAATGTCAGAGTGGCCATCTTTGTTGTAGCAAC 311  
Qy 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100  
Db 312 TGTGGCCCAAGCTACATGTTGTCCAACTTGGCGGGGCCCTTTGGGATCCATTGCAAC 371  
Qy 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
Db 372 TTGGCTATCGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAATAATATGCTCTTCTGGA 431  
Qy 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140  
Db 432 TGTGAATAACTCTGCCACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTAGG 491  
Qy 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrPheGlnGlySerLeuAspAlaVal 160  
Db 492 CCTTATTCTCTCGTCCGCTCGTGTCTTCTGTAAATGGCAGGCTCTCTGGATGCTGTA 551  
Qy 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180  
Db 552 ATGCCCCCATCTGATGCATCAGCATAGTCCATTACAAACCTACAGGAGAGGATATAGTT 611  
Qy 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetMetGlnSerCys 200  
Db 612 NNTCTTGCTACAGACATTAACTTCTCTGTGTCTGTGACTGGGTGATGATGAGTCTCTGT 671  
Qy 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAspGlyHisGlnGln 220  
Db 672 TTTGGCTTTTCACTTTTCATGCTAGTCTAGAGAANCAGAAAAATACGATGTCACACAGCAG 731  
Qy 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsn 236  
Db 732 TTCTTTTCGCATCGTACAGCTGATAGGAACACGTCAGCAAGCTGAAAAAT 779

Search completed: April 25, 2005, 09:25:45  
Job time : 3698 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 25, 2005, 06:45:00 ; Search time 204 Seconds  
(without alignments)  
2390.247 Million cell updates/sec

Title: US-10-679-246-2

Perfect score: 1611

Sequence: 1 MWIIIFLLPPYVFISEMSRQ.....IAQLFAENGUNGINVTISM 298

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO spoel/US10679246/runat\_22042005.185511.7719/app\_query.fasta\_1.455  
-DB=Issued Patents NA -QPMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10679246 @CGN 1.1.69 @runat\_22042005.185511.7719 -NCPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq:  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq:  
5: /cgn2\_6/prodata/1/ina/PTCUS\_COMB.seq:  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	1274	4	US-09-591-694-1
2	1611	100.0	2440	4	Sequence 588, App
3	1489	92.4	1884	4	Sequence 11, Appl
4	447	27.7	1420	3	Sequence 1, Appl
5	436	26.4	257	4	Sequence 397, App
6	336	20.2	1036	3	Sequence 39, Appl
7	325	20.2	1425	3	Sequence 200, App
8	217	13.5	8050	3	Sequence 11, Appl
9	217	13.5	8050	3	Sequence 11, Appl
10	130.5	8.1	2327	4	Sequence 1378, A
11	128.5	8.0	2308	4	Sequence 29760, A
12	121	7.5	1463	4	Sequence 23, Appl

13	112.5	7.0	2854	2	US-08-724-394A-17	Sequence 17, Appl
14	111	6.9	1260	4	US-09-949-016-1949	Sequence 1949, Ap
15	108.5	6.7	2248	1	US-08-639-237-1	Sequence 1, Appl
16	108.5	6.7	2248	1	US-08-975-405-1	Sequence 6, Appl
17	108.5	6.7	2264	3	US-09-167-109-6	Sequence 6, Appl
18	108	6.7	1029	4	US-09-270-767-13746	Sequence 13746, A
19	107.5	6.7	563	3	US-09-325-932A-40	Sequence 40, Appl
20	106	6.6	1335	4	US-09-774-528-33	Sequence 33, Appl
21	104.5	6.5	1926	3	US-09-249-588A-2	Sequence 2, Appl
22	104.5	6.5	1926	3	US-09-410-399-3	Sequence 3, Appl
23	104.5	6.5	2580	3	US-09-050-863-2	Sequence 2, Appl
24	104.5	6.5	2580	3	US-09-359-081-2	Sequence 2, Appl
25	104.5	6.5	2943	3	US-09-221-017B-153	Sequence 153, App
26	104.5	6.5	5452	2	US-09-130-114-1	Sequence 1, Appl
27	104.5	6.5	8705	4	US-09-647-344A-14	Sequence 14, Appl
28	104.5	6.5	9600	3	US-08-910-647-1	Sequence 1, Appl
29	104.5	6.5	9600	3	US-09-620-925-1	Sequence 1, Appl
30	104.5	6.5	10596	1	US-07-884-811-15	Sequence 15, Appl
31	104.5	6.5	10596	1	US-07-885-971-15	Sequence 15, Appl
32	104.5	6.5	10596	1	US-08-087-783A-15	Sequence 15, Appl
33	104.5	6.5	10596	1	US-08-194-088B-15	Sequence 15, Appl
34	104.5	6.5	10596	2	US-08-194-087-15	Sequence 15, Appl
35	104.5	6.5	10596	5	PCT-US93-04648-15	Sequence 15, Appl
36	104.5	6.5	16080	4	US-09-724-566A-48	Sequence 48, Appl
37	104.5	6.5	16080	4	US-09-471-669A-48	Sequence 48, Appl
38	103.5	6.4	1329	4	US-09-270-767-13968	Sequence 13968, A
39	103.5	6.4	8572	4	US-09-949-016-13691	Sequence 13691, A
40	103	6.4	1023	4	US-09-902-540-3824	Sequence 3824, Ap
41	103	6.4	13807	3	US-09-052-469-5	Sequence 5, Appl
42	103	6.4	13807	4	US-08-422-582-5	Sequence 5, Appl
43	103	6.4	13807	4	US-09-052-262-5	Sequence 5, Appl
44	103	6.4	13941	4	US-09-799-451-341	Sequence 341, App
45	103	6.4	14060	3	US-08-658-136-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-591-694-1

; Sequence 1, Application US/09591694

; Patent No. 6638734

; GENERAL INFORMATION:

; APPLICANT: John C. Reed

; APPLICANT: Shu-ichi Matsuzawa

; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved

; FILE REFERENCE: P-LJ 4220

; CURRENT APPLICATION NUMBER: US/09/591,694

; CURRENT FILING DATE: 2000-06-09

; EARLIER APPLICATION NUMBER: US 09/330,517

; EARLIER FILING DATE: 1999-06-11

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 1274

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (274)...(1167)

US-09-591-694-1

Alignment Scores:

Pred. No.: 1.45e-171 Length: 1274  
Score: 1611.00 Matches: 298  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-10-679-246-2 (1-298) x US-09-591-694-1 (1-1274)

Qy 1 MetValIleIlePheLeuLeuProPofTyValPheIleSerCluMetSerArgGln 20

```

RESULT 2
US-09-023-655-588
; Sequence 588, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE

```

QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTrpGlnGlySerLeuAspAlaVal 160  
DB 927 CCTATTCTCTGTCGGTCCCTGCTGCTCTCTGTAATGGCAAGCTCTCTGGATGCTGA 986  
QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180  
DB 987 ATGCCCATCTGATGATCAGCATAGTCCATTTACACCTTACAGGAGAGGATATAGTT 1046  
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSerCys 200  
DB 1047 TTTCTTGCTACAGACATTAATCTCTCTGCTGCTGTTGACTGGGTGATGATGCACTCTGT 1106  
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlnGlyTrpAspGlyHisGlnGln 220  
DB 1107 TTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAATAACGATGTCACCAAGCAG 1166  
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240  
DB 1167 TTTCTTGGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1226  
QY 241 LeuGluLeuAsnGlyHisArgArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260  
DB 1227 CTTGAGCTAAATGCTCATAGGCGCAGATTGACTTTGGGAAGCGACTCTCGATCTATTCTAT 1286  
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280  
DB 1287 GAAGGAATTCACACGCCATTATGAATAGCGACTGTCTAGTCTTTGACACCGCATTGCA 1346  
QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298  
DB 1347 CAGCTTTTGGCAGAAATGGCAATTTAGGCATCAATGTAATCTATTCTTCCATGTGT 1400

## RESULT 3

US-09-544-618-11  
; Sequence 11, Application US/09544618  
; Patent No. 6503502  
; GENERAL INFORMATION:  
; APPLICANT: Telerman, Adam  
; APPLICANT: Amson, Robert  
; APPLICANT: Cohen, Daniel  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, PROTEINS, DRUGS AND DIAGNOSTIC  
; FILE OF INVENTION: AGENTS OF USE IN TREATING CANCER  
; FILE REFERENCE: 065691-0139  
; CURRENT APPLICATION NUMBER: US/09/544,618  
; CURRENT FILING DATE: 2000-03-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 1884  
; TYPE: DNA  
; ORGANISM: TSAP 3  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(846)  
US-09-544-618-11

## Alignment Scores:

Pred. No.:	1.56e-157	Length:	1884
Score:	1489.00	Matches:	280
Percent Similarity:	99.29%	Conservative:	1
Best Local Similarity:	98.94%	Mismatches:	1
Query Match:	92.43%	Indels:	2
DB:	4	Gaps:	0

US-10-679-246-2 (1-298) x US-09-544-618-11 (1-1884)

QY 17 MetSerArgGlnThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSerGln 36  
DB 1 ATGAGCGGTGACACTGCTACAGCATTAACCTACCGGTACCTCGAAGTGTCCACCATCCAG 60  
QY 37 ArgValProAlaLeuThrGlyThrAlaSerAsnAspLeuAlaSerLeuPheGlu 56  
DB 61 AGGGTGCCCTGCTGCTGACACACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 120

QY 57 CysProValCysPheAspTyrValLeuProIleLeuGlnCysGlnSerGlyHisLeu 76  
DB 121 TGTCCAGTCTGCTTTGACTATGTGTACGCCCATTTCTTCAATGTGAGAGTGCCATCTT 180  
QY 77 ValCysSerAsnCysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGly 96  
DB 181 GTTTGTAGCACTGTGCGCCCAAGCTCAGATGTGTCCAACTTTGCGGGGCCCTTTGGGA 240  
QY 97 SerIleArgAsnLeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyr 116  
DB 241 TCCATTGCACTGCTGCTATGGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATAT 300  
QY 117 AlaSerSerGlyCysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeu 136  
DB 301 GCGTCTCTCTGGATGTGAAATAAATCTCTGCCACACACAGAAAAAGCAGACCATGAAGAGCTC 360  
QY 137 CysGluPheAspProTyrSerCysProCysProGlyAlaSerCysLysTrpGlnGlySer 156  
DB 361 TGTGAGTTTGTAGGCTTTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 157 LeuAspAlaValMetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGly 176  
DB 421 CTGGATGCTGTAATGCCCATCTGATGATCAGCATAAGTCCATTACACCTTACAGGGA 480  
QY 177 GluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMet 196  
DB 481 GAGGATATAGTTTCTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 197 MetGlnSerCysPheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAsp 216  
DB 541 ATGCACTGCTGTTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAAATACGAT 600  
QY 217 GlyHisGlnGlnPhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsn 236  
DB 601 GGTCAACAGCAGTCTTCTGCAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 660  
QY 237 PheAlaTyrArgLeuGluLeuAsnGlyHisArgArgArgLeuThrTrpGluAlaThrPro 256  
DB 661 TTTGCTTACGACTTGAGCTAAATGCTATAGCGCAGCATTTGACTTTGGGAAGCAGCTCTCT 720  
QY 257 ArgSerIleHisGluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAsp 276  
DB 721 CGATCTATTATGAAAGAAATTCGACAGCCATTTATGAATAGCGACTGTCTAGTCTTTGAC 780  
QY 277 ThrSerIleAlaGlnLeuPheAlaGlu-AsnGlyAsnLeuGlyIleAsnValThrIleSe 296  
DB 781 CC-AGCATTCACAGCTTTTTCGACAAATGGCAATTTAGGCATCAATGTAATCTATTCTTTC 839  
QY 296 rMetCys 298  
DB 840 CATGTGT 846  
RESULT 4  
US-09-362-506-1  
; Sequence 1, Application US/09362506  
; Patent No. 6111167  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Maize SINA Orthologue-1 and Uses Thereof  
; FILE REFERENCE: 0936  
; CURRENT APPLICATION NUMBER: US/09/362,506  
; CURRENT FILING DATE: 1999-07-27  
; EARLIER APPLICATION NUMBER: 60/100,258  
; EARLIER FILING DATE: 1998-09-14  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1420  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: CDS



```

; LOCATION: (202)... (1140)
US-09-362-506-1
Alignment Scores:
Pred. No.: 3,12e-40 Length: 1420
Score: 447.00 Matches: 103
Percent Similarity: 53.65% Conservative: 44
Best Local Similarity: 37.59% Mismatches: 101
Query Match: 27.75% Indels: 26
DB: 9 Gaps: 9

US-10-679-246-2 (1-298) x US-09-362-506-1 (1-1420)
QY 39 ProAlaLeuThrGlyThrAlaSerAsnAspLeuAlaSer----- 53
DB 316 CCGTGTCTACATCGGCATCGCTAGCAATGTTGCAATGTCGTCTGTGCTTGAAT 375
QY 54 ---LeuPheGluCysProValCysPheAspTyrValLeuProPheLeuGlnCysGln 72
DB 376 GATTGCTCGAATGCCAGTGTGTACCAACTCGATCGCCCACTTACTTCCAGTGGCCA 435
QY 73 SerGlyHisLeuValCysSerAsnCysArgProLysLeu---ThrCysCysProThrCys 91
DB 436 AATGGCCACACGATCTCTAGTTGCAAGCAGGCTAGAGAACCAATGCCCAACCTGT 495
QY 92 ArgGlyProLeuGlySerLeuArgAsnLeuAlaMetGluLysValAlaAsnSerValLeu 111
DB 496 CGCCAGAACTGGAAACATCATAGTCTCTCGAAGGTGGCAGAGCACTCCAG 555
QY 112 PheProCysLysTyrAlaSerSerGlyCysGluLeuThrLeuProHisThrGluLysAla 131
DB 556 CTTCCGTGCAAGTACAGAGCAGCGGATGCCAGATCCACCCATACAGCAAACTG 615
QY 132 AspHisGluGluLeuCysGluPheArgProTyrSerCysProCysProGlyAlaSerCys 151
DB 616 AAGCAGCAGAGCTCTGCAGGTTTCAGCGCTCAGCTGCCGTACGACGTTCCGAGTGC 675
QY 152 LysTrpGlnGlySerLeuAspAlaValMetProHisLeuMetHisGlnHisLysSerIle 171
DB 676 CTGATCCAGCGGACGCTCCGCTTCTGTTGTTCCATCTCTATCAACCACTATAG---GTG 732
QY 172 ThrThrLeuGlnGlyGluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAla 191
DB 733 GACTTGCACAGGGCTGCACGTTCAACCCCGGTACGTGAGGCCCAAC---CGTACGAG 789
QY 192 ValAsp-----TrpValMet---MetGlnSerCysPheGlyPheHisPheMetLeu 207
DB 790 GTGGAGAAATGCCAGTGGATGCTCACTGTGTTTCAAGTGTTCGGGACGACTTCTGCCCTG 849
QY 208 -----ValLeuGluLysGlnGluLysTyrAspGlyHisGlnGlnPhePhe 222
DB 850 CACTTCAGCGGCTGCTGCTGGGATGGCGCGGTGTAC-----ATG 891
QY 223 AlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArgLeuGlu 242
DB 892 GCGTTCCTCGGTTTCATGGCGAGGAGAGCGAGCGCGGCGGTTCGGGTACAGCCTGGAG 951
QY 243 LeuAsnGlyHisArgArgGluLeuThrTrpGluAlaThrProArgSerIleHisGluGly 262
DB 952 GTGGCGCGGCGCGCGGAGAGTACGTGGCAGGCGACGCGCGGAGCGGTACGAGGACAGC 1011
QY 263 IleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAlaGlnLeu 282
DB 1012 CACCGGAAGTGGCGGACAGCTTCAGCGGCTCATCTATCCACAGAACATGGGGCTTTC 1071
QY 283 PheAlaGluAsnGly-----AsnLeuGlyIleAsnValThr 294
DB 1072 TTCTCGGCGCGCGGACGAGGAGTCAAGTGGCGGTCAAC 1113

RESULT 5
US-09-016-434-397
; Sequence 397, Application US/09016434
; Patent No. 6500938

GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 397:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT03
CLONE: 2113436
US-09-016-434-397

Alignment Scores:
Pred. No.: 4,58e-39 Length: 257
Score: 426.00 Matches: 73
Percent Similarity: 92.94% Conservative: 6
Best Local Similarity: 85.88% Mismatches: 6
Query Match: 26.44% Indels: 0
DB: 4 Gaps: 0

US-10-679-246-2 (1-298) x US-09-016-434-397 (1-257)
QY 120 GlyCysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPhe 139
DB 3 GCGTGTCTCTCCCTGACCTCCATACCGAAGAACCGAACATGACACATATGTGAATAC 62
QY 140 ArgProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAla 159
DB 63 CGTCCCTACTCTCTGCCCCATGCTCTGCTGCTCTCTGCAAGTGGCAGGGGTCCCTGGAAGCT 122
QY 160 ValMetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIle 179
DB 123 GTGATGTCCCATCTCATGCCGCCCAAGAGGATTACCACTTCCAGGAGGAGACATC 182
QY 180 ValPheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSer 199
DB 183 GTCTTTCTAGCTACACACATTAATCTTCCAGGGGCTGTCTGACTGGGTGATGTCAGTCA 242
QY 200 CysPheGlyPheHis 204
DB 243 TGTTTTGGCCATCAC 257
```

```

RESULT 6
US-09-325-932A-39
; Sequence 39, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of forestry plant develop
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 1036
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-39

Alignment Scores:
Pred. No.: 8-27e-27 Length: 1036
Score: 325.00 Matches: 90
Percent Similarity: 46.48% Conservative: 42
Best Local Similarity: 31.69% Mismatches: 114
Query Match: 20.24% Indels: 38
DB: 3 Gaps: 9

US-10-679-246-2 (1-298) x US-09-325-932A-39 (1-1036)
QY 36 GlnArgValProAlaLeuThrGly-----ThrThrAlaSerAsnAsnAsp 50
Db 175 CGCAGATTTCTGCTGTTAGGGATAAAACCGAGCAGGTTGCGAGCGCATCATTTGGGGAT 234
QY 51 Leu-----AlaSerLeuPheGluCysProValCysPheAspTyrValLeuProProlle 68
Db 235 TTGGATGGAGCTTCACTAATAGAACTCTCT---GGCAGAGAAACCCCTCCAAATTTGTATA 291
QY 69 LeuGlnCysGlnSerGlyHisLeuValCysSerAsnCysArgProLysLeuThr---Cys 87
Db 292 ACAGAGTGCCTGCTGCTAGCATATGTTCTGGATGTAATCCAGGTTAAGCACTGG 351
QY 88 CysProThrCysArgGlyProLeuGlySerIleArgAsnLeuAlaMetGluLysValala 107
Db 352 TGTCCTGCTGCGAGATGCAACCTTGGAAATTTTAGGTGCTTAGCTCCTGAAACGGAGACA 411
QY 108 AsnSerValLeuPheProCysLysTyrAlaSerSerGlyCysGluLeuThrLeuProHis 127
Db 412 TCATCTCAAGAACTTACTTTCATGATATCAAAAGCTATGTTGTGAGGATATGTATCCTTAC 471
QY 128 ThrGluLysAlaAspHisGluGluLeuValCysGluPheArgProTyrSerCysProCysPro 147
Db 472 TACAGTGAATTAAGACATGAAGCTAGTGCATGATTTAGGCCATACAACTGCTCCTATGCT 531
QY 148 GlyAlaSerCysLysTyrGlnGlySerLeuAspAlaValMetProHisLeuMetHisGln 167
Db 532 GGCTCCGAATGCAAGCTAGTTGGAGATATTCCTTTTGGTGCTCATTTAAGAGATGAT 591
QY 168 HisLys-----SerIleThrThr 173
Db 592 CACAAAGTTTATATGCATAATAGTTGCACCTTTGCATCATGATGTAAGTCAAAATCCA 651
QY 174 LeuGlnGlyGluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAlaValasp 193
Db 652 CTCGAGTTGAGAATGCTATTTTGGATGCCAACCTGTAATCAAT-----693
QY 194 TrpValMetMetGlnSerCysPheGlyPheHisPheMetLeuValLeuGluLysGlnGlu 213
Db 694 -----TGTTTTGGCAATTTCTTTGCTACATTTTGA---GGCTTT 732
QY 214 LysTyrAspGlyHisGlnGlnPhePheAlaIleValGlnLeuIleGlyThrArgLysGln 233
Db 732 -----

US-09-325-932A-200
; Sequence 200, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of forestry plant develop
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 200
; LENGTH: 1425
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-200

Alignment Scores:
Pred. No.: 1-8e-26 Length: 1425
Score: 325.00 Matches: 90
Percent Similarity: 46.48% Conservative: 42
Best Local Similarity: 31.69% Mismatches: 114
Query Match: 20.17% Indels: 38
DB: 3 Gaps: 9

US-10-679-246-2 (1-298) x US-09-325-932A-200 (1-1425)
QY 36 GlnArgValProAlaLeuThrGly-----ThrThrAlaSerAsnAsnAsp 50
Db 171 CGCAGATTTCTGCTGTTAGGGATAAAACCGAGCAGGTTGCGAGCGCATCATTTGGGGAT 230
QY 51 Leu-----AlaSerLeuPheGluCysProValCysPheAspTyrValLeuProProlle 68
Db 231 TTGGATGGAGCTTCACTAATAGAACTCTCT---GGCAGAGAAACCCCTCCAAATTTGTATA 287
QY 69 LeuGlnCysGlnSerGlyHisLeuValCysSerAsnCysArgProLysLeuThr---Cys 87
Db 288 ACAGAGTGCCTGCTGCTAGCATATGTTCTGGATGTAATCCAGGTTAAGCACTGG 347
QY 88 CysProThrCysArgGlyProLeuGlySerIleArgAsnLeuAlaMetGluLysValala 107
Db 348 TGTCCTGCTGCGAGATGCAACCTTGGAAATTTTAGGTGCTTAGCTCCTGAAACGGAGACA 407
QY 108 AsnSerValLeuPheProCysLysTyrAlaSerSerGlyCysGluLeuThrLeuProHis 127
Db 408 TCATCTCAAGAACTTACTTTCATGATATCAAAAGCTATGTTGTGAGGATATGTATCCTTAC 467
QY 128 ThrGluLysAlaAspHisGluGluLeuValCysGluPheArgProTyrSerCysProCysPro 147
Db 468 TACAGTGAATTAAGACATGAAGCTCCTCAATTTTAGGCCATACAACTGCTCCTATGCT 527
QY 148 GlyAlaSerCysLysTyrGlnGlySerLeuAspAlaValMetProHisLeuMetHisGln 167
Db 527 -----

```





QY	239	TyrArgLeu-----	241
Db	2280	CACAAGCTTTTATGAGAGCGCCGCTTCTCGGCTTCAATCAGCAGTGGTGTGAAG	2339
QY	242	-----GluLeuAsn-----	252
Db	2340	GCGCGCATCAACAACAGCAGCGCGATCGGATCAGTCAACAGCAGCGACCATCACTAT	2399
QY	253	Glu-----AlaThrProArgSerIleHis	260
Db	2400	CATCTGATCTCAAGACCAAGACAGCAGCGCCATGAGCATACAT	2444
RESULT 11			
US-09-270-767-29760			
; Sequence 29760, Application US/09270767			
; Patent No. 6703491			
; GENERAL INFORMATION:			
; APPLICANT: Homburger et al.			
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster			
; FILE REFERENCE: File Reference: 7326-094			
; CURRENT APPLICATION NUMBER: US/09/270,767			
; NUMBER OF SEQ ID NOS: 62517			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 29760			
; LENGTH: 2308			
; TYPE: DNA			
; ORGANISM: Drosophila melanogaster			
US-09-270-767-29760			
Alignment Scores:			
Pred. No.:	0.000564	Length:	2308
Score:	128.50	Matches:	54
Percent Similarity:	34.16%	Conservative:	29
Best Local Similarity:	72.22%	Mismatches:	72
Query Match:	7.98%	Indels:	89
DB:	4	Gaps:	12
US-10-679-246-2 (1-298) x US-09-270-767-29760 (1-2308)			
QY	24	AlaLeuProThrGlyThrSerIysCys---ProProSerGlnArgVal-----	38
Db	1663	GCATGCCAGATGGCCATCTGA-TGTGCGCGCTGTTTCAGGCATCTGCTGGCGGATG	1721
QY	39	-----ProAlaLeuThrGlyThrThrAlaSerAsn	48
Db	1722	GACGGTATGAACCCACACTCTAGAAACTAACTAAGATCCTTGCAAAATGCCACCACCAAC	1781
QY	49	AsnAspLeuAlaSerLeuPheGluCysProValCysPheAsp-----	62
Db	1782	TCCAACTAACGAGTTCGTTTACTGTCTCT-----TACGACCTTTTITTTTCAATA	1835
QY	63	TyrValLeuProPheLeuGlnCysGlnSerGlyHisLeuValCysSerAsnCysArg	82
Db	1836	TATTCCTCTCTAGCTTG-----CGC	1856
QY	83	ProIysLeuThrCysCysProThrCysArgGlyProLeuGly-----SerIleArg	99
Db	1857	GATCAGATTGCCAGTCGCCCAATTCGCGGTGGAAATTTCTAAGACCATGCGTGGCGC	1916
QY	100	AsnLeuAlaMetGluIysValAlaAsnSerValLeuPheProCysIysTyrAlaSerSer	119
Db	1917	AACITGGCGGTGGAGAGCGGCTCAGAAATGCGCAGCGAGTGTCAATTC-----	1967
QY	120	GlyCysGluIleThrLeuProHisThrGluIysAlaAspHisGluGlu-----LeuCysGlu	138
Db	1968	---TGCAACAAGAGATTCCCATACAAATCTCTCGAACGCCATGAACAAACAGCAGTGGCCAG	2024
QY	139	PheArgProTyrSerCysProCysProGlyAlaSerCysIysTyrPheGlnGlySerLeuAsp	158
Db	2025	GAGCGCCCGACCAAGTGCAATAATACCATCGCATTCGGTGTTCAGTGGCGGGAGACCTTACCAC	2084

QY	159	AlaValMetProHis-----LeuMetHisGlnHisIysSerIleThrThrLeuGln	175
Db	2085	GAGACCAACGAGCATGAGCGCACTGCTTGCAATCCCGAAGTCT-----	2129
QY	176	GlyGluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrVal	195
Db	2129	-----	2129
QY	196	MetMetGlnSerCysPheGlyPheHisPheMetLeuValLeuGlu-----	210
Db	2130	-----GGCTACGAGTAATGCGAGCCCTGGAGGCCCGACGATAGG	2171
QY	211	---LysGlnGluIysTyrAspGlyHisGlnGlnPhePheAlaIleValGlnLeuIleGly	229
Db	2172	ATCAAGGAGGAGAAG-----AAGATGTTTAAACACCTGATTGCTTGTAGC	2219
QY	230	ThrArgIys	232
Db	2220	TACGAGAAG	2228
RESULT 12			
US-09-620-312D-23			
; Sequence 23, Application US/09620312D			
; Patent No. 6569662			
; GENERAL INFORMATION:			
; APPLICANT: Tang, Y. Tom			
; APPLICANT: Liu, Chenghua			
; APPLICANT: Asundi, Vinod			
; APPLICANT: Zhang, Jie			
; APPLICANT: Ren, Feiyan			
; APPLICANT: Chen, Rui-hong			
; APPLICANT: Zhao, Qing A.			
; APPLICANT: Wehrman, Tom			
; APPLICANT: Xue, Aidong J.			
; APPLICANT: Yang, Yonghong			
; APPLICANT: Wang, Jian-Rui			
; APPLICANT: Zhou, Ping			
; APPLICANT: Ma, Yungqing			
; APPLICANT: Wang, Dunrui			
; APPLICANT: Wang, Zhiwei			
; APPLICANT: John Tillinghast			
; APPLICANT: Drmanac, Radoje T.			
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and			
; TITLE OF INVENTION: Polypeptides			
; FILE REFERENCE: 784CIP2B			
; CURRENT APPLICATION NUMBER: US/09/620,312D			
; CURRENT FILING DATE: 2000-07-19			
; PRIOR APPLICATION NUMBER: 09/552,317			
; PRIOR FILING DATE: 2000-04-25			
; PRIOR APPLICATION NUMBER: 09/488,725			
; PRIOR FILING DATE: 2000-01-21			
; NUMBER OF SEQ ID NOS: 1105			
; SOFTWARE: pt_FL_genes Version 1.0			
; SEQ ID NO 23			
; LENGTH: 1463			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (178) .. (1002)			
US-09-620-312D-23			
Alignment Scores:			
Pred. No.:	0.00189	Length:	1463
Score:	121.00	Matches:	30
Percent Similarity:	39.64%	Conservative:	14
Best Local Similarity:	27.03%	Mismatches:	49
Query Match:	7.51%	Indels:	18
DB:	4	Gaps:	4
US-10-679-246-2 (1-298) x US-09-620-312D-23 (1-1463)			
QY	67	ProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsnCys-----	81

Db 148 CCTCTTCGAGTGTACTAATGTGTCACCTTGATGTGGCTGCTGTTTATCCACCTACTA 207  
QY 82 -----ArgProLysLeuThrCysCysProThrCysArgGlyProLeuGly 96  
Db 208 GCAGATGCCGGCTGAAGGAGGAGCAGGCCACGTGCCCAATTGCTGTTGAGATCAGT 267  
QY 97 Ser-----IleArgAsnLeuAlaMetGluLysValAlaAsnSerValLeuPhePro 113  
Db 268 AAGAGCCTCTGCTGCCGGAACCTGGCGTGGAGAAAGCCGTGACGCGAGCTGCCCTTCAGAG 327  
QY 114 CysLysTyrAlaSerSerGlyCysGluIleThrLeuProHisThrGluLysAlaAspHis 133  
Db 328 TGTGGCTTC-----TGCCTGGCCAGTTTCCCGCTCCCTCTCTGGAGAGGCAC 375  
QY 134 Glu---GluLeuCysGluPheArgProTyrSerCysProCysProGlyAlaSerCysLys 152  
Db 376 CAGAAAGAGGAGTGCAGGAGCAGGGTAACCCAGTGAAGTACAAACGCATCGCGCTGCCCA 435  
QY 153 TrpGlnGlySerLeuAspAlaValMetProHis 163  
Db 436 TGGCAGGGCCCTTCCATGAGCTGACGGTGCAC 468

## RESULT 13

US-08-724-394A-17

; Sequence 17, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies Thereeto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2854 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..2854

; OTHER INFORMATION: /note= "cDNA 27"

; US-08-724-394A-17

Alignment Scores:  
Pred. No.: 0.0507 Length: 2854  
Score: 112.50 Matches: 57  
Percent Similarity: 32.46% Conservative: 30  
Best Local Similarity: 21.27% Mismatches: 96  
Query Match: 6.98% Indels: 85  
DB: 2 Gaps: 12

US-10-679-246-2 (1-298) x US-08-724-394A-17 (1-2854)

QY 37 ArgValProAlaLeuThrGlyThrAlaSerAsnAspLeuAlaSerLeuPheGlu 56  
Db 7 CGCGTCCGAAAGCTATGGCTTCAACCCAGCACCAAGAGATGATGGAGGAGCCACC 66  
QY 57 CysProValCysPheAspTyrValLeuProProIle----LeuGlnCysGlnSerGlyHis 75  
Db 67 TGTCTCATCTGCTGAGCCTGATGACCAACCCAGTAAGCATCACTGT-----GGACAC 120  
QY 76 LeuValCysSerAsnCys-----Arg 82  
Db 121 AGCTACTGCCACTTGTGTATTAACAGACTTCTTTAAAAACCAAGCCAAAAAGCACTGAGG 180  
QY 83 ProLysLeuThrCysCysProThrCysArgGlyPro-----LeuGlySerIleArgAsn 100  
Db 181 CAGGAGACATTCTGCTGCTCCCGAGTGTGGGCTCCATTTCATATGATAGCTCCGA--- 237  
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120  
Db 238 -----CCNACAGCAGCTGGGAAGCCTC 261  
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGlu----- 138  
Db 262 ATTGAAGCC-----CTCAAAGAGACGGATCAAGAAATGTCTGTGAGGAACAC 309  
QY 139 -----PheArgProTyrSerCysProCysProGlyAlaSerCysLysTyrTrpGlnGly 155  
Db 310 GGAGAGCAGTTCCACCTGTC-----TCGGAAGACGAGGGG 345  
QY 156 SerLeuAspAlaValMetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGln 175  
Db 346 CAGCTCATCTGCTGGCGCTGTGAGCGGGCCACACACACACACACACAGCTCTT 405  
QY 176 GlyGluAspIleValPheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpVal 195  
Db 406 GTTGAAGACGTA----- 417  
QY 196 MetMetGlnSerCysPheGlyPheHisPheMetLeuValLeuGluLysGlnLysTyr 215  
Db 418 -----TGCAGGGCTACAG-----GAAAAGCTCCAGAAA--- 447  
QY 216 AspGlyHisGlnGlnPhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGlu 235  
Db 448 -----GCTGTGACAAACTG-----AAGCAACTTGAA 474  
QY 236 AsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThr 255  
Db 475 GACAGATGTACGAGCAGAGCTGTCCACAGCAATCGGAATTAATAATGGAAGAAGAAG 534  
QY 256 ProArgSerIleHisGluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPhe 275  
Db 535 GTACAGATTTCAGAGACAAAAATCCGCTGTGACTTTAAGAATCTCCAGTGTTCCTACAT 594  
QY 276 AspThrSerIleAlaGlnLeuPhe 283  
Db 595 GAGGAAGAGAGTCTTATCTCTGG 618

## RESULT 14

US-09-949-016-1949

; Sequence 1949, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1949
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1949

Alignment Scores:
Pred. No.:          Length:          1260
Score:             0.0199           111.00
Percent Similarity: 42.86%           50
Best Local Similarity: 35.71%         53
Query Match:       6.89%             29
DB:                8
Gaps:              4

US-10-679-246-2 (1-298) x US-09-949-016-1949 (1-1260)
QY 24 AlaLeuProThrGlyThrSerLysCysPro---ProSerGlnArg-----Val 38
Db 751 TCCCTGCTGCTGGTACCTGAGCTGCCACACCCAGGAGAGAGGTCAGTCTGAAAGTG 810
QY 39 ProAlaLeuThrGlyThrAlaSerAsnAsnAspLeuAlaSerLeuPheGlu----- 56
Db 811 CCCAGGAG-CCAGGAGCCAGGAGTGTGGAGGGCGACGCTGCGCGCTGCAGGAGGAGG 869
QY 57 ---CysProValCysPheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHis 75
Db 870 ACGTGCAAGTGTGCTTGCAGCCGCGCGTG---TCCATGTGCTTTGTGCGCGGCCAC 926
QY 76 LeuValCysSerAsnCysArgProLysLeuThrCysCysProThrCysArgGlyProLeu 95
Db 927 CTGCTGTGCTGAGTGTGCGCGCGCTGCAGCTGTGCGCCCATCTGCAGAGCCCGCTC 986
QY 96 GlySerIleArgAsnLeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLys 115
Db 987 CGCAGCGG-CGT-----GGCACCTTCTGCTAGGCCAGGTGCCATGGCCGCCGCG 1039
QY 116 TyrAlaSer-SerGlyCysGluIleThrLeuProHisThrGluLysAlaAspHisGlu 135
Db 1040 TGGGCTGCAGATGGG-----CTCCCT-----GCCCG 1066
QY 135 uLeuCysGluPheArgProTyrSerCysProCysProGlyAlaSerCysLysTrpGln 154
Db 1067 TCTCTGCTGTTTCTGACTGTGTTCTGGGCTGCTGAGGA-----TGGCAG 1112

RESULT 15
US-08-639-237-1
; Sequence 1, Application US/08639237
; Patent No. 5710013
; GENERAL INFORMATION:
; APPLICANT: Xiong, Jessie
; APPLICANT: Goeddel, David V.
; TITLE OF INVENTION: No. 5710013el Protein - TRAF6
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,237
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J
; REGISTRATION NUMBER: 24,7747
; REFERENCE/DOCKET NUMBER: T96-004/A63312
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2248 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 230..1795
; US-08-639-237-1

Alignment Scores:
Pred. No.:          Length:          2248
Score:             0.0973           108.50
Percent Similarity: 31.40%           48
Best Local Similarity: 19.81%        128
Query Match:       6.73%             157
DB:                1
Gaps:              19

US-10-679-246-2 (1-298) x US-08-639-237-1 (1-2248)
QY 5 1lePheLeuLeuProTyrValPheIleSerGluMetSerArgGlnThrAlaThrAla 24
Db 176 CTCCTGTTGGGCTCTCCCGGCGCACTAGAACGAGCAGTGAATCAAGTTACTATGAGT 235
QY 25 LeuProThrGlyThrSerLysCysProSerGln----- 36
Db 236 CTGCTAAACTGTGAAACAGAGCTGTGGATCCAGCCAGTCTGAAAGTGAAGTGTGTCGCC 295
QY 37 -----ArgValProAlaLeuThr-----GlyThrThrAlaSerAsn 48
Db 296 ATGGCCAGCTCTCTAGCGCTCTAACAAGAGATGATGTGGTGGAACTGCCAGCAGC 355
QY 49 AsnAspLeuAlaSerLeuPhe----- 55
Db 356 GGGAACTCTCCAGCTCATTTATGGAGGAGATCCAGGAGATATGATGTAGAGTTGACCCA 415
QY 56 -----GluCysProValCysPheAspTyrValLeuProProlleLeu 69
Db 416 CCCCTGAAAGCAAGTATGAATGCCCATCTCTCTG---ATGCGATTACGAGAGCAGTG 472
QY 70 GlnCysGlnSerGlyHisLeuValCysSerAsnCysArgProLysLeuThrCysCysPro 89
Db 473 CAAAGCCCATCGGCCCATAGGTTCTGCAAGAGCTGCATCATATAA----- 517
QY 90 ThrCysArgGlyProLeuGlySerIleArgAsn-----LeuAlaMet 103
Db 518 -----TCAATAGGGATGAGGTCACAAATGTCAGTTGACAT 556
QY 104 GluLysValAlaAsnSerValLeuPheProCysLysTyrAla----- 117
Db 557 GAAATACTCTGGAAATCACTATTTCCAGACAATTTTGCAAAACCTGAGATCTTTCT 616
QY 118 -----SerSerGlyCys-----GluIleThrLeuProHisThrGlu 129
Db 617 CTGATGTTGAAATGTCCAAATGAAGTTGTTTGCACAGATGGAAGTGCATCTTGAG 676

```



```
OY 130 LysAlaAspHisGluLeuCysGluPhe-----Arg 140
Db 677 -----GATCATCAAGCACATTGTGAGTTTGTCTTATGGATTGTCCCAATGCCAGCGT 730
OY 141 ProTyr-----142
Db 731 CCCTTCCAAAANTTCATATTATATTCACATTCTGAAGGATTGTCCAAGGAGACAGGTT 790
OY 143 -----Ser 143
Db 791 TCTTGTGACAACTGTGTCATCAATGGCATTGAAGATAAAGAGATCCATGACCAGAAC 850
OY 144 CysProCysProGlyAlaSerCysLysTrpGlnGlySerLeuAspAlaValMetProHis 163
Db 851 TGTCTCTTGGCAAAATGTCATCTGTGAATACTGCAATACTATA-----CTCATCGAGAA 904
OY 164 LeuMetHisGlnHisLysSerIle-ThrThrLeuGlnGlyGluAspIleValPheLeuAl 183
Db 905 CAGATGCCTAATCATATATGATCTAGACTGCCCTACAGCCCAA-----TTCCATGC 955
OY 183 aThrAspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSerCysPheGlyPh 203
Db 956 ACATTCAGTACTTTTG-----GTTGCCATGAAAAGATGCAGAGGAT 997
OY 203 eHisPheMetLeuValLeuGlnLysGlnGlyLysTyrAspGlyHisGlnGlnPheAl 223
Db 998 -CAC-----TTGGCAGCGCCACTACAGAGAAACACCCAGTCACACATGAGAATGTTGGC 1050
OY 223 a-----IleValGlnLeuIleGlyThrArgLysGlnAlaGluAs 236
Db 1051 CCAGGCTGTTTCATAGTTTGAGCGTTATACCCGACTCTGGGTATATCTCAGAGGTCGGAA 1110
OY 236 nPheAlaTyr-----ArgLeuGluLeuAsnGlyHisArgArg 249
Db 1111 TTTCAGGAAACTATTTCACGAGTTAGAGGTCGCTTGTGAAGACAAGACCATCAATCCG 1170
OY 249 gLeuThrTrpGluAlaThrProArgSerIleHisGluGlyIleAlaThrAlaIleMetAs 269
Db 1171 G-----GAGCTGACTGCTAAAAATGGAACCTCAGAGTATGTATGTAAGTGAGCTCAA 1221
OY 269 nSerAspCysLeuValPheAspThrSerIleAlaGlnLeuPheAlaGluAsn-----286
Db 1222 ACGAACCATTCGAACCCCTTTGAGGACAAAGTTGCTGAAATCGAAGCACACAGCTGCAATGG 1281
OY 287 -----GlyAsnLeuGlyIleAsnVal 293
Db 1282 AATTTATATTGGAGATTGGCAACTTTGGAATGCATTG 1321
```

Search completed: April 25, 2005, 09:29:18  
Job time : 219 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: April 25, 2005, 08:26:55 ; Search time 654 Seconds  
(without alignments)  
2766.320 Million cell updates/sec

Title: US-10-679-246-2  
Perfect score: 1611  
Sequence: 1 MWILFLPPYFISEMSRQ.....IAQLFAENGLNINTWMC 298

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5633728 seqs, 3035525691 residues  
Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10679246/runat\_22042005\_185512\_7796/app\_query.fasta\_1.455  
-DB=Published Applications NA -QWMT=fastac -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -MMAP=US10679246 @CGN 1.1 480 @runat\_22042005\_185512\_7796  
-NCPU=6 -ICPU=3 -NO WPAR -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBIOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubna/US10\_NEW\_PUB.seq:\*  
20: /cgn2\_6/ptodata/2/pubna/US11\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/2/pubna/US60\_NEW\_PUB.seq:\*  
22: /cgn2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1611	100.0	1274	18	US-10-679-246-1	Sequence 1, Appli
2	1611	100.0	2440	17	US-10-641-643-588	Sequence 588, App
3	1611	100.0	2924	18	US-10-357-930-22959	Sequence 22959, A
4	1611	100.0	2924	18	US-10-357-930-25457	Sequence 25457, A
5	1611	100.0	2924	18	US-10-357-930-28822	Sequence 28822, A
6	1611	100.0	6107	10	US-09-764-891-8070	Sequence 8070, Ap
7	1603	99.5	2829	17	US-10-108-260A-269	Sequence 269, App
8	1206	74.9	2502	15	US-10-177-293-426	Sequence 426, App
9	1199	74.4	975	17	US-10-172-118-1117	Sequence 1117, Ap
10	1199	74.4	975	17	US-10-342-887-1117	Sequence 1117, Ap
11	1198	74.4	2240	15	US-10-177-293-424	Sequence 424, App
12	1198	74.4	2240	18	US-10-717-597-137	Sequence 137, App
13	1198	74.4	2240	18	US-10-755-889-123	Sequence 123, App
14	474	29.4	1535	9	US-09-925-297-84	Sequence 84, Appli
15	462.5	28.7	1438	18	US-10-437-963-83416	Sequence 83416, A
16	447	27.7	1408	17	US-10-425-114-2756	Sequence 2756, Ap
17	446	27.7	1522	18	US-10-425-115-182305	Sequence 182305,
18	446	27.7	360	9	US-09-960-352-10011	Sequence 10011, A
19	429.5	26.7	1444	17	US-10-425-114-4151	Sequence 4151, Ap
20	429.5	26.7	1667	18	US-10-425-115-28950	Sequence 28950, A
21	429	26.6	984	9	US-09-938-842A-463	Sequence 463, App
22	429	26.6	984	11	US-09-938-842A-463	Sequence 463, App
23	426	26.4	257	17	US-10-305-720-397	Sequence 397, App
24	422	26.2	1448	18	US-10-437-963-84963	Sequence 84963, A
25	419	26.0	1398	17	US-10-424-599-98299	Sequence 98299, A
26	417.5	25.9	1383	18	US-10-324-120-4	Sequence 4, Appli
27	416	25.8	1526	17	US-10-424-599-21498	Sequence 21498, A
28	416	25.8	1683	18	US-10-425-115-27188	Sequence 27188, A
29	415	25.8	2542	18	US-10-437-963-9355	Sequence 9355, A
30	413.5	25.7	1540	17	US-10-424-599-18853	Sequence 18853, A
31	410.5	25.5	856	18	US-10-767-795-1843	Sequence 1843, Ap
32	405	25.1	1388	18	US-10-425-115-61385	Sequence 61385, A
33	405	25.1	1835	18	US-10-437-963-55834	Sequence 55834, A
34	404	25.1	1378	17	US-10-424-599-15841	Sequence 15841, A
35	398.5	24.7	1638	18	US-10-437-963-29521	Sequence 29521, A
36	397	24.6	1740	17	US-10-424-599-18852	Sequence 18852, A
37	358	22.2	2078	17	US-10-424-599-21496	Sequence 21496, A
38	338	21.0	1342	18	US-10-425-115-183094	Sequence 183094,
39	332	20.6	777	18	US-10-767-701-3143	Sequence 3143, Ap
40	326	20.2	921	18	US-10-425-115-147280	Sequence 147280,
41	326	20.2	1036	14	US-10-219-220-39	Sequence 39, Appli
42	326	20.2	1319	14	US-10-219-220-228	Sequence 228, App
43	325	20.2	1425	14	US-10-219-220-200	Sequence 200, App
44	324.5	20.1	832	18	US-10-425-115-32733	Sequence 32733, A
45	319.5	19.8	2142	17	US-10-108-260A-1336	Sequence 1336, Ap

ALIGNMENTS

RESULT 1  
US-10-679-246-1  
; Sequence 1, Application US/10679246  
; Publication No. US20040163138A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Matsuzawa, Shu-ichi  
; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved  
; TITLE OF INVENTION: in Protein Degradation, Products and Methods Related Thereto  
; FILE REFERENCE: 66821-235  
; CURRENT APPLICATION NUMBER: US/10/679,246  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: US 09/591,694  
; PRIOR FILING DATE: 2000-06-09  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1274  
; TYPE: DNA  
; ORGANISM: Homo sapien



```
QY 61 PheAspTyrValLeuProPheLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80
Db 687 TTGACTATGTGTACCGCCCATCTTCAATGTCAGAGTGGCCATCTGTGTTGTAGCAAC 746
QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
Db 747 TGTGCGCCAAAGCTCACATGTTGTCCAACTTGCCTGGGGGCCCTTTGGGATCCATTCGCAC 806
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120
Db 807 TTGGCTATGAGAAAGTGCGTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGGA 866
QY 121 CysGluLeuThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140
Db 867 TGTGAATAAATCTCTGCCACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTTGG 926
QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal 160
Db 927 CCTTATCTCTGTCGCTGGTGTCTTCTGTAAATGCGAAGGCTCTCTGGATGCTGTA 986
QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180
Db 987 ATGCCCCCATCTGATGCATCAGCATAGTCCATTTACAACCTACAGGAGGAGATAGTT 1046
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetGlnSerCys 200
Db 1047 TTCTTGTGTACAGACATTAATCTTCTGTGTCTGTGTGCTGTGATGATGATGCTCTGT 1106
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAspGlyHisGlnGln 220
Db 1107 TTTGGCTTTCACTTCATGTAGTCTTAGAGAAACAGGAAAATACGATGTCACGACGAG 1166
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240
Db 1167 TTCTTCGCAATCGTACAGCTGATAGGAACAGCAAGCAAGCTGAAATTTTGTCTACCGA 1226
QY 241 LeuGluLeuAsnGlyHisArgArgArgLeuThrTyrGluAlaThrProArgSerIleHis 260
Db 1227 CTTGAGCTAAATGGTCATAGGCGACGATTTGACTTGGGAAGCGACTCTCTCATTTTCAT 1286
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280
Db 1287 GAAGGAATTTGCAACGCCATTTAATGATAGGACTGCTAGTCTTTGACACCGACATTCGA 1346
QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298
Db 1347 CAGCTTTTGCAGAAATGCGCAATTTAGGCATCAATGTAACATTTTCCATGTGT 1400

RESULT 3
US-10-357-930-22959
; Sequence 22959, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007ECN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
```

```
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22959
; LENGTH: 2924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..57..2923..2924
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-22959

Alignment Scores:
Pred. No.: 4..69e-189 Length: 2924
Score: 1611.00 Matches: 298
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-679-246-2 (1-298) x US-10-357-930-22959 (1-2924)

QY 1 MetValIleIleIlePheLeuLeuProProTyrValPheIleSerGluMetSerArgGln 20
Db 753 ATGGTTATAATTAATTTTCTCTCTGCTCTCTTATGTATTTATTTTCAGAAATGAGCCGTCAG 812
QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSerGlnArgValProAla 40
Db 813 ACTGCTACAGCATTAACCTACCGGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTGCC 872
QY 41 LeuThrGlyThrThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCys 60
Db 873 CTGACTGGCACAACCTGCATCCAACTGCTTGGCGAGTCTTTTTCAGTGTCCAGTCTGC 932
QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80
Db 933 TTTGACTATGTCTTACCGCCCATCTTCAATGTCAGAGTGGCCATCTTCTTTGTAGCAAC 992
QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
Db 993 TGTGCCCAAGCTCACATGTTGTCCAACTTCCCGGGGCCCTTTGGGATCCCATTCGCAAC 1052
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120
Db 1053 TTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGGA 1112
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluLeuCysGluPheArg 140
Db 1113 TGTGAAATAAATCTGCCACACAGAAAAAGCAGACCATGAAGAGCTCTGTGAGTTTAGG 1172
QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal 160
Db 1173 CTTTATTTCTGTGCTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1232
QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180
Db 1233 ATGCCCATCTGATGCATCAGCATTAATTCATTAACCTACAGGAGGAGATAGTT 1292
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetGlnSerCys 200
Db 1293 TTTCTTGTCTACAGACATTAATCTTCTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 1352
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAspGlyHisGlnGln 220
Db 1353 TTTGGCTTTCACTTCATGTAGTCTTTAGAGAAACAGGAAAAATACGATGTCACGACGAG 1412
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240
Db 1413 TTCTTCGCAATCGTACGCTGATAGAACACGCAAGCAGCTGAAATTTTCTCTTACCAG 1472
```

```
QY 241 LeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260
Db 1473 CTTGAGCTAAATGGTCATAGGCGAGGATTGACTTGGGAAGGACCTCTCGATCTATTCA 1532
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280
Db 1533 GAAGGAATTCGCAACGCCATTATGAATAGCGACTGTCTAGTCTTTGACACCGAGCATTCGA 1592
QY 281 GinLeuPheAlaGluAsnGlyHisLeuGlyIleAsnValThrIleSerMetCys 298
Db 1593 CAGCTTTTTCAGAAAATGGCAATTTAGGCATCAATGTAATCTTTCCATGTTGT 1646

RESULT 4
US-10-357-930-25457
; Sequence 25457, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25457
; LENGTH: 2924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 57, 2923, 2924
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-25457

Alignment Scores:
Pred. No.: 4,69e-189 Length: 2924
Score: 1611.00 Matches: 298
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-679-246-2 (1-298) x US-10-357-930-25457 (1-2924)
QY 1 MetValIleIlePheLeuLeuProProTyThrValPheIleSerGluMetSerArgGln 20
Db 753 ATGGTTATTAATTTCTTCCTGCTCTCTTATGATTTATTTCAGAAATGAGCGGTGAG 812
QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProProSerGlnArgValProAla 40
Db 813 ACTGCTACGATTAATCTACCGGTACCTCGAAGTGTCCACCATCCAGAGGGTGCCTGCC 872
QY 41 LeuThrGlyThrThrAlaSerAsnAsnAspLeuAlaSerIlePheGluCysProValCys 60
Db 873 CTGACTGGCACAACTGCATCCCAATGACTTGGCGAGTCTTTTGGAGTGTCCAGTCTGC 932
```

```
QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80
Db 933 TTTGACTATGTTTACCGCCCATCTTCAATGTCAGAGTGGCCATCTTGTGTTGTAGCAAC 992
QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
Db 993 TGTCTCCCAAGCTCACATGTTGTCCAACTTGCCTGGGGCCCTTTGGGATCCATTGCGCAAC 1052
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120
Db 1053 TGGGTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTAATATATGCGTCTTCTGGA 1112
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140
Db 1113 TGTGAATTAATCTCTCCACACACAGAAAAGCAGACCATGAAGAGCTCTGTGAGTTAGG 1172
QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerIleAspAlaVal 160
Db 1173 CCTTATTCCTGTCCGTGCTGCTTCTTCTGTAATGCGAAGGCTCTCTGGATGCTGTA 1232
QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180
Db 1233 ATGCCCATCTGATGATCAGCATTAAGTCCATTACACCCCTACAGGAGAGGATATAGTT 1292
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSerCys 200
Db 1293 TTTCTTGTCTACAGACATTAATCTTCTGCTGCTGTTGACTGGTGATGATGCGCTCTGT 1352
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGluLysTyrAspGlyHisGlnGln 220
Db 1353 TTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGAGAAAATACGATGGTCCACGAC 1412
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240
Db 1413 TTTCTTGGCAATCGTACGCTGATAGAAACACCAAGCAAGCTGAAAATTTTCTTACCGA 1472
QY 241 LeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260
Db 1473 CTTGAGCTAAATGGTCATAGGCGAGGATTGACTTGGGAAGGAGCTCTCGATCTATTCA 1532
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280
Db 1533 GAAGGAATTCGCAACGCCATTATGAATAGCGACTGTCTAGTCTTTGACACCGAGCATTCGA 1592
QY 281 GinLeuPheAlaGluAsnGlyHisLeuGlyIleAsnValThrIleSerMetCys 298
Db 1593 CAGCTTTTTCAGAAAATGGCAATTTAGGCATCAATGTAATCTTTTCCATGTTGT 1646

RESULT 5
US-10-357-930-28822
; Sequence 28822, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
```

```
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28822
; LENGTH: 2924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..57..2923, 2924
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28822

Alignment Scores:
Pred. No.: 4,69e-189 Length: 2924
Score: 1611.00 Matches: 298
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-10-679-246-2 (1-298) x US-10-357-930-28822 (1-2924)

QY 1 MetValIleIlePheLeuLeuProProTyrValPheIleSerGluMetSerArgGln 20
|
|
|
Db 753 ATGGTTATAATTATTTCTCTGCTCCTCTATGTATTTATTTTCAGAAATGAGCCGTCAG 812

QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla 40
|
|
|
Db 813 ACTGCTACAGCATTAACCTACCGTACCTCGAAGTGTCACCATCCAGAGGGTGCCTGCC 872

QY 41 LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys 60
|
|
|
Db 873 CTGACTGGCAACAACGATCCCAACATGACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGC 932

QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80
|
|
|
Db 933 TTTGACTATGTGTACCGCCCATTTCTCAATGTCAGAGTGGCCATCTGTGTTGTAGCAAC 992

QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
|
|
|
Db 993 TGTGCGCCCAAGCTACATGTTGTGTCCAACTGTCGCGGGGCCCTTTTGGGATCCATTCGCAAC 1052

QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120
|
|
|
Db 1053 TTGGCTATGGAGAAAGTGAGTAACTTTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGGA 1112

QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluLeuCysGluPheArg 140
|
|
|
Db 1113 TGTGAAATAACTCTGCCACACACAGAAAAAGCAGACCATGAGAGCTCTGTGAGTTTAGG 1172

QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal 160
|
|
|
Db 1173 CCTTATTCTGTGCGGCCCTGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1232

QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrLeuGlnGlyGluAspIleVal 180
|
|
|
Db 1233 ATGCCCCATCTGATGATCAGCATAGTCCATTTACAACTTACAGGAGGAGGATATAGTT 1292

QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetMetGlnSerCys 200
|
|
|
Db 1293 TTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1352

QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlyTyrAspGlyHisGlnGln 220
|
|
|
Db 1353 TTTGGCTTTTCACTTCATGTAGTCTTAGAGAAACAGGAAAAATACGATGCTCACCAGCAG 1412

QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240
|
|
|
Db 1413 TTTCTTGGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAAATTTTGTCTTACCGA 1472
```

```
QY 241 LeuGluLeuAsnGlyHisArgArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260
|
|
|
Db 1473 CTTGAGCTAAATGGTTCATAGGCGACGATGACTTGGGAAAGCGACTCTCTCGATCTATTTCAT 1532

QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280
|
|
|
Db 1533 GAAGGAATTGCAACAGCCATTATGAATAGGACTGTCTAGTCTTTGACACACCATGCA 1592

QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298
|
|
|
Db 1593 CAGCTTTTTCAGAAAAATGGCAATTTAGGCATCAATGTAACATTTTCCATGTCT 1646

RESULT 6
US-09-764-891-8070
; Sequence 8070, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8070
; LENGTH: 6107
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-8070

Alignment Scores:
Pred. No.: 1.4e-188 Length: 6107
Score: 1611.00 Matches: 298
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-679-246-2 (1-298) x US-09-764-891-8070 (1-6107)
```

```
QY 1 MetValIleIlePheLeuLeuProProTyrValPheIleSerGluMetSerArgGln 20
|
|
|
Db 3437 ATGGTTATAATTATTTCTCTGCTCCTTATGTATTTATTTTCAGAAAAATGAGCCGTCAG 3496

QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla 40
|
|
|
Db 3497 ACTGCTACAGCATTAACCTACCGTACCTCGAAGTGTCACCATCCAGAGGGTGCCTGCC 3556

QY 41 LeuThrGlyThrThrAlaSerAsnAspLeuAlaSerLeuPheGluCysProValCys 60
|
|
|
Db 3557 CTGACTGGCACAACGTCATCCCAACATGACTTGGCGAGTCTTTTTCAGTGTCCAGTCTGC 3616

QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80
|
|
|
Db 3617 TTTGACTATGTGTACCGCCCATTTCTCAATGTCAGAGTGGCCATCTTCTGTTGTAGCAAC 3676

QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
|
|
|
Db 3677 TGTGCGCCCAAGCTCACATGTTGTCCAACTTCCGCGGGGCCCTTTGGGATCCATTCGCAAC 3736

QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120
|
|
|
Db 3737 TTGGCTATGGAGAAAGTGCTTAATTCAGTACTTTTCCCTGTAAATATGCGTCTTCTGGA 3796

QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluLeuCysGluPheArg 140
|
|
|
Db 3797 TGTGAAATAACTCTGCCACACACAGAAAAAGCAGACCATGAAAGAGCTCTGTGAGTTTAGG 3856

QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal 160
|
|
|
Db 3857 CCTTATTCTGTGCGGCCCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3916
```



```
QY 161 MetProHisLeuMethHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180
Db 3917 ATGCCCAATCTGATCATCAGCATAGTTCATACAACTTACAGGAGGATAGTT 3976
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetValMetGlnSerCys 200
Db 3977 TTTCCTTGCTACAGACATTAATCTTCCTGGTGTCTTGAAGTGGTGTATGATGCGTCTCTGT 4036
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAspGlyHisGlnGln 220
Db 4037 TTGGCTTTTCACCTTCATGTTAGTCTTAGAGAAACAGGAAATATACGATGGTACCAGCAG 4096
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240
Db 4097 TTCTTCGCATCTGATCAGCTGATAGAACACGACAGCAGCTGAAAATTTTCTTACCAG 4156
QY 241 LeuGluLeuAsnGlyHisArgArgArgLeuThrTyrGluAlaThrProArgSerIleHis 260
Db 4157 CTTGAGCTAAATGGTTCATAGGCGCAGTATGAAATAGCAGTCTTGGGAGCGACTCCTCGATCTATTCAT 4216
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280
Db 4217 GAAGGAATTTGCAACAGCCATATGAAATAGCAGTCTTGGGAGCGACTCCTCGATCTATTCAT 4276
QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValIleSerMetCys 298
Db 4277 CAGCTTTTTCGAGAAATGCAATTTAGGCATCAATGTAATTTCCATGTGT 4330
RESULT 7
US-10-108-260A-269
; Sequence 269, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 269
; LENGTH: 2829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-269
Alignment Scores:
Pred. No.: 4,41e-188 Length: 2829
Score: 1603.00 Matches: 297
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 1
Query Match: 99.50% Indels: 0
DB: 17 Gaps: 0
US-10-679-246-2 (1-298) x US-10-108-260A-269 (1-2829)
QY 1 MetValIleIlePheLeuProProTyrValPheIleSerGluMetSerArgGln 20
Db 1407 ATGGTTATATATATTTCTCTCGCTCTCTTATGTTATTTTTCAGAAATGAGCGGTCTAG 1466
QY 21 ThrAlaThrAlaLeuProThrGlyThrSerLysCysProSerGlnArgValProAla 40
Db 1467 ACTGCTACAGCATTAACCTACCGGTACTCTGAAAGTGTCCACCATCCAGAGGGTGGCTGCC 1526
QY 41 LeuThrGlyThrThrAlaSerAsnAsnLeuAlaSerLeuPheGluCysProValCys 60
Db 1527 CTGACTGGCACAACCTGCATCCCAATAGCTTGGCGAGTCTTTTTCAGTGTCCAGTCTCC 1586
QY 61 PheAspTyrValLeuProProIleLeuGlnCysGlnSerGlyHisLeuValCysSerAsn 80
Db 1587 TTGTACTATGTATCCGCCCAATCTTCAATGTCAGAGTGGCCATCTTGTGTGTAGCAAC 1646
```

```
QY 81 CysArgProLysLeuThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
Db 1647 TGTTCGCCCAAGACTCACATGTTGTCCAATTCGCGGGGCCCTTTGGGATCCATTCGCAAC 1706
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120
Db 1707 TTGGCTATGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTAAATATATGCGTCTTCTGGA 1766
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140
Db 1767 TGTGAAATAACTCTCCACACACAGAAAAGACAGACCACTGAAGACTCTGTGAGTTTAGG 1826
QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrGlnGlySerLeuAspAlaVal 160
Db 1827 CCTTATCTCTGTCGCTGCGCTGCTTCTTCTGTAATGCAAGGCTCTCTGCGATGCTGA 1886
QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180
Db 1887 ATGCCCACTCTGATCATCAGCATTAAGTCCATTACACCCCTACAGGAGAGGATATAGTT 1946
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTyrValMetGlnSerCys 200
Db 1947 TTTCCTGCTACAGCATTAATCTTCTGCTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCT 2006
QY 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAspGlyHisGlnGln 220
Db 2007 TTTCGCTTTTCACTTCTCATGTTAGTCTTTAGAGAAACAGGAAATATACGATGTCACGACG 2066
QY 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240
Db 2067 TTCTTCGCAATCGTACAGCTGATAGGAACACGCAAGCAAGCTGAAAATTTTTCCTTACCGA 2126
QY 241 LeuGluLeuAsnGlyHisArgArgArgLeuThrTyrGluAlaThrProArgSerIleHis 260
Db 2127 CTTGAGCTAAATGCTCATAGGCGAGCATTTGACTTGGGAGCGACTCTCTGATCTATTCAT 2186
QY 261 GluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAla 280
Db 2187 GAAGGAATTTGCAACAGCCATTAATGATAGCGACTCTCTAGTCTTTTGGCCACGACATTCGA 2246
QY 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298
Db 2247 CAGCTTTTTCGAGAAATGCGCAATTTAGGCATCAATGTAATTTTCCATGTGT 2300
RESULT 8
US-10-177-293-426
; Sequence 426, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
```

```
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 426
; LENGTH: 2502
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-426

Alignment Scores:
Pred. No.:      8,66e-139      Length:      2502
Score:          1206.00        Matches:    228
Percent Similarity: 78.18%      Conservative: 30
Best Local Similarity: 69.09%    Mismatches: 30
Query Match:      74.86%       Indels:      42
DB:               15          Gaps:         5

US-10-679-246-2 (1-298) x US-10-177-293-426 (1-2502)

QY 10 ProTyrValPheIleSerGluMetSerArgGlnThrAlaThrAlaLeuProThrGlyThr 29
Db 506 CCATGGTTGTTGGAGCGCGATGAGCCCGCGTCTCCACCGGC---CCAGCGCTAAT 562

QY 30 SerLysCys-----
Db 563 AAACCTGCAGCAAGCAGCGCGCGCCAGCACACACCTCCGTCGCCGGTGGCGCC 622

QY 33 -----ProSerGlnArgValProLa----- 40
Db 623 CGCGCGCGCCACCATCTCGGTGCGGGCGCGCGTCTCGCGGTGCGCGCGCGCG 682

QY 41 -----LeuThrGly-----ThrThrAlaSerAsn 49
Db 683 GCGGTGATCTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 742

QY 50 AspLeuAlaSerLeuPheGluCysProValCysPheAspTyrValLeuProIleLeu 69
Db 743 GAGCTGACCTCGCTCTTCGAGTGTCCGGTCTGCTTTGACTATGTCTGCTCTATTCTG 802

QY 70 GlnCysGlnSerGlyHisLeuValCysSerAsnCysArgProLysLeuThrCysCysPro 89
Db 803 CAGTGCAGCGCGCGCACCTGGTGTGTAAACCAATGCCGCCAGAGTTGAGCTGCTGCCG 862

QY 90 ThrCysArgGlyProLeu---GlySerIleArgAsnLeuAlaMetGluLysValAlaAsn 108
Db 863 ACGTGCAGGGCGCGCTGACGCCACCATCAGAACTGGCTATGGAAGTGGCGCTCG 922

QY 109 SerValLeuPheProCysLysTyrAlaSerSerGlyCysGluIleThrLeuProHisThr 128
Db 923 GCAGTCTGTTCCTGTAGTATGTCACACCGGGTGTTCCTGACCTTGACCATACG 982

QY 129 GluLysAlaAspHisGluLysGluPheArgProTyrSerCysProCysProGly 148
Db 983 GAGAACCAAGACATCAGACATATGTAATACCGTCTACTCTGCGCATGCTCTGGT 1042

QY 149 AlaSerCysLysTrpGlnGlySerLeuAspAlaValMetProHisLeuMetHisGlnHis 168
Db 1043 GCTTCTCGAAGTGGCAGGGTCCCTGGAAGCTGTGATGTCCATCTCATGACGCCAC 1102

QY 169 LysSerIleThrThrLeuGlnGlyGluAspIleValPheLeuAlaThrAspIleAsnLeu 188
Db 1103 AAGAGCATTAACACCCCTTCAGGGAGAAGACATCGTCTTCTAGCTACAGACATTAAC 1162
```

```
QY 189 ProGlyAlaValAspTrpValMetMetGlnSerCysPheGlyPheHisPheMetLeuVal 208
Db 1163 CCAGGGGCTGTGCACTGGGTGATGATGAGTCAGTCATGTTTGGCCATCCTCATGCTGGG 1222

QY 209 LeuGluLysGlnGluLysTyrAspGlyHisGlnGlnPhePheAlaIleValGlnLeuIle 228
Db 1223 CTGGAGAAACAAGAGAGTACGAGGCCACACAGAGTCTTTTGGCCATCGTCTGCTCAT 1282

QY 229 GlyThrArgLysGlnAlaGluAsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArg 248
Db 1283 GGCACCCCGCAGCAAGCCGAGAACTTTGCTCAGACTGGAGTTGAATGGGAACCGCGCG 1342

QY 249 ArgLeuThrTrpGluAlaThrProArgSerIleHisGluGlyIleAlaThrAlaIleMet 268
Db 1343 AGATTGACCTGGGAGGCCACGCCCGTTCGATTCATGACGGTGTGCTCGGCCATCATG 1402

QY 269 AsnSerAspCysLeuValPheAspThrSerIleAlaGlnLeuPheAlaGluAsnGlyAsn 288
Db 1403 AACAGCGACTGCTGTGTTTTCGACACAGCCATAGCACATCTTTTGCAGATAATGGGAC 1462

QY 289 LeuGlyIleAsnValThrIleSerMetCys 298
Db 1463 CTTGGAATCAATGTTACTATTCTACATGT 1492

RESULT 9
US-10-172-118-1117
; Sequence 1117, Application US/10172118
; Publication NO. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1117
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_005067
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1117

Alignment Scores:
Pred. No.:      1,58e-138      Length:      975
Score:          1199.00        Matches:    228
Percent Similarity: 84.90%      Conservative: 25
Best Local Similarity: 76.51%    Mismatches: 26
Query Match:      74.43%       Indels:      20
DB:              17          Gaps:         5

US-10-679-246-2 (1-298) x US-10-172-118-1117 (1-975)

QY 9 ProProTyrValPheIleSerGluMetSerArgGlnThrAlaThrAlaLeuProThrGly 28
Db 107 CCACCA-----TCTCGGCTCGGGCCCCCGCTCGTCG----- 139

QY 29 ThrSerLysCysProProSerGln-----ArgValProAlaLeuThrGly----- 43
Db 140 -----CGGTGCGCGCGCGCGGTGATCTCGGGCCCCCGCGCGCGC-GGCGGGGC 192

QY 44 -----ThrThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCysPhe 61
Db 44 -----TCTCGGCTCGGGCCCCCGCTCGTCG----- 139
```

```
Db 193 GGCCCGGTGTCGCCGAGCACCAGAGCTGACCTCGCTCTTCGAGTGTCCGGTCTGCTTT 252
Qy 62 AspTyrValLeuProProlleuLeuGlnCysGlnSerArgGlnThrAlaThrAlaLeuProThrGly 81
Db 253 GACTATGCTCGCTCCTATTCTGCAGTCCAGGCCGACCTGTGTGTATACCAATGC 312
Qy 82 ArgProLysLeuThrCysCysProThrCysArgGlyProLeu---GlySerIleArgAsn 100
Db 313 CGCCAGAGTTGAGCTGCTGCCGAGCTGCAGGGCGCCCTGACGCCAGCATCAGGAAC 372
Qy 101 LeuAlaMetGluLysValAlaAlaSerValLeuPheProCysLysTyrAlaSerSerGly 120
Db 373 CTGGCTATGAGAGAGTGGCTCGGCAGTCTCTTCCCTGTAAGTATGCCACCGGC 432
Qy 121 CysGluIleThrLeuProHisThrGlnLysAlaAspHisGluLysLeuCysGluPheArg 140
Db 433 TGTTCCTGACCTGACCATACGGAGAACACAGAACATGATGTAATACCGT 492
Qy 141 ProTyrSerCysProCysProGlyAlaSerCysLysTrpGlnGlySerLeuAspAlaVal 160
Db 493 CCTACTCTGCCCCATGCTCTGCTTCTCAAGTGGCAGGGGTCCCTGGAAGCTGTG 552
Qy 161 MetProHisLeuMetHisGlnHisLysSerIleThrLeuGlnGlyGluAspIleVal 180
Db 553 ATGTCCCATCTGACCATACGGAGAACACAGAACATGATGTAATACCGT 492
Qy 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetGlnSerCys 200
Db 613 TTTCAGCTACAGACATTAACCTTGCAGGGGTGTCGACTGGGTGATGATGATGATG 672
Qy 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAspGlyHisGlnGln 220
Db 673 TTTCAGCTACATCTCATGCTGCTGCTGGAGAACAGAGAGATGACAGAGCCACCG 732
Qy 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240
Db 733 TTTTTCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
Qy 241 LeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260
Db 793 CTGGAGTTGAATGGGAACCGCGGAGATGACCTGGAGGCCACGCCCTTCGATTCAT 852
Qy 261 GluGlyIleAlaThrAlaIleMetAsnSerCysLeuValPheAspThrSerIleAla 280
Db 853 GACGGTGTGCTGCGGCCCATCATGACAGCGAGTGCCTGTTTTCGACACAGCCATGCA 912
Qy 281 GlnLeuPheAlaGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys 298
Db 913 CATCTTTTTCAGATAATGGGAACCTTGGAAATCAATGTTACTATTCTACATGT 966
```

## RESULT 10

```
US-10-342-887-1117
; Sequence 1117, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
```

```
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1117
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1117
Alignment Scores:
Pred. No.: 1,58e-138 Length: 975
Score: 1199.00 Matches: 228
Percent Similarity: 84.90% Conservative: 25
Best Local Similarity: 76.51% Mismatches: 26
Query Match: 74.43% Indels: 20
DB: 17 Gaps: 5
```

US-10-679-246-2 (1-298) x US-10-342-887-1117 (1-975)

```
Qy 9 ProProTyrValPheIleSerGluMetSerArgGlnThrAlaThrAlaLeuProThrGly 28
Db 107 CCACCA-----TCTGGCTGCGGGCCCGGCTCGTCCG----- 139
Qy 29 ThrSerLysCysProProSerGln-----ArgValProAlaLeuThrGly----- 43
Db 140 -----CGTGGCCCGCGCGCGGTGATCTCGGGCCCGCGCGCGGC-GGCGGGGCC 192
Qy 44 -----ThrThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCysPhe 61
Db 193 GGCCCGGTGTCGCCGAGCACCAGAGCTGACCTGCTTTCGAGTGTCCGGTCTGCTTT 252
Qy 62 AspTyrValLeuProProlleuLeuGlnCysGlnSerGlyHisLeuValCysSerAsnCys 81
Db 253 GACTATGCTCGCTCCTATTCTGCAGTGGCAGGCCGACCTGCTGTGTAAACCAATGC 312
Qy 82 ArgProLysLeuThrCysCysProThrCysArgGlyProLeu---GlySerIleArgAsn 100
Db 313 CGCCAGAGTTGAGCTGCTGCCGAGCTGCAGGGCGCCCTGACGCCAGCATCAGGAAC 372
Qy 101 LeuAlaMetGluLysValAlaAlaSerValLeuPheProCysLysTyrAlaSerSerGly 120
Db 373 CTGGCTATGAGAGAGTGGCTCGGCAGTCTCTTCCCTGTAAGTATGCCACCGGC 432
Qy 121 CysGluIleThrLeuProHisThrGlnLysAlaAspHisGluLysLeuCysGluPheArg 140
Db 433 TGTTCCTGACCTGACCATACGGAGAACACAGAACATGATGTAATACCGT 492
Qy 141 ProTyrSerCysProCysProGlyAlaSerCysLysTrpGlnGlySerLeuAspAlaVal 160
Db 493 CCTACTCTGCCCCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552
Qy 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180
Db 553 ATGTCCCATCTGACCATACGGAGAACACAGAACATGATGTAATACCGT 612
Qy 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAspTrpValMetGlnSerCys 200
Db 613 TTTCAGCTACAGACATTAACCTTGCAGGGGTGTCGACTGGGTGATGATGATGATG 672
Qy 201 PheGlyPheHisPheMetLeuValLeuGluLysGlnGlyLysTyrAspGlyHisGlnGln 220
Db 673 TTTCAGCTACATCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
Qy 221 PhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTyrArg 240
Db 733 TTTTTCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
Qy 241 LeuGluLeuAsnGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHis 260
Db 793 CTGGAGTTGAATGGGAACCGCGGAGATGACCTGGAGGCCACGCCCTTCGATTCAT 852
Qy 261 GluGlyIleAlaThrAlaIleMetAsnSerCysLeuValPheAspThrSerIleAla 280
Db 853 GACGGTGTGCTGCGGCCCATCATGACAGCGAGTGCCTGTTTTCGACACAGCCATGCA 912
```

[illegible]



Qy	90	ThrCysargGlyProLeu--GlySerIleArgAsnLeuAlaMetGluLysValAlaAsn	108
Db	863	ACGTGCAGGGCGCCCTCACGCCCCAGCATACGAACCTGGCTATAGGAGAAGGTGGCCCTCG	922
Qy	109	SerValLeuPheProCysLysTyTyrAlaSerSergLyCysGluIleThrLeuProHisThr	128
Db	923	GCAGTCCTGTTTCCCTGTAACTATGCACCACAGGGGCTGTTCCCTGTGACCCCTGCACCAATACG	982
Qy	129	GluLysAlaaspHisGluGluLeuCysGluPheArgProTyrSerCysProCysProGly	148
Db	983	GAGAAAACAGAACATGAAGACAATATGCAATACCGCCCTACTCTCTCCCATGTCTTGCTGGT	1042
Qy	149	AlaSerCysLysTrpGlnGlySerLeuAspAlaValMetProHisLeuMetHisGlnHis	168
Db	1043	GCTTCCTGCAAGTGCAGGGGTCCCTGGAAGCTGTGATGTCCCATTCTCATGTGACGCCCCAC	1102
Qy	169	LysSerIleThrThrLeuGlnGlyGluAspIleValPheLeuAlaThrAspIleAsnLeu	188
Db	1103	AAGAGCATTAACACCCCTTCAGGAAGAAGACATCGTCTTCTTAGCTACAGACATAAATTC	1162
Qy	189	ProGlyAlaValAspTrpValMetMetGlnSerCysPheGlyPheHisPheMetLeuVal	208
Db	1163	CCAGGGGCTGTGCACTGGGTGATGATGCAATGCTATGTTTTGGCCATCACTTCATGCTGGTG	1222
Qy	209	LeuGluLysGlnGluLysTyrAspGlyHisGlnGlnPhePheAlaIleValGlnLeuIle	228
Db	1223	CTGAGAAAAACAAGAGAAGTACGAAGGCACACAGCAGTTTTTTGGCCATCGCTCTGCTCATT	1282
Qy	229	GlyThrArgLysGlnAlaGluAsnPheAlaTyrArgLeuGluLeuAsnGlyHisArgArg	248
Db	1283	GGCACCCGCAAGCAAGCCGAGAATTGGCTACAGACTGGAGTTGAATGGAAACCGCGCG	1342
Qy	249	ArgLeuThrTrpGluAlaThrProArgSerIleHisGluGlyIleAlaThrAlaIleMet	268
Db	1343	AGATTGACCTGGGAGGCCACGCCCGCTCGAATTCATGACGGTGTGGCTGGCGGCATCATG	1402
Qy	269	AsnSerAspCysLeuValPheAspThrSerIlealagLnLeupheAlaGluAsnGlyAsn	288
Db	1403	AACAGCACTGCCCTGTGTTTTTCGACACAGCAATAGCACATCTTTTTCAGATATATGGAAAC	1462
Qy	289	LeuGlyIleAsnValThrIleSerMetCys	298
Db	1463	CTTTGGAAATCAATGTTACTATTTCTACATGT	1492

## RESULT 14

US-09-925-297-84  
; Sequence 84, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 84  
; LENGTH: 1535  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-297-84

Alignment Scores:	
Pred. No.:	4.52e-48
Score:	474.00
Length:	1535
Percent Similarity:	98.95%
Matches:	94
Best Local Similarity:	98.95%
Conservative:	0
Query Match:	29.42%
Mismatches:	0
Indels:	1

DB:	9	Gaps:	0
US-10-679-246-2 (1-298) x	US-09-925-297-84 (1-1535)		
Qy	205	PheMetLeuValLeuGluLeuGlyGlnGluLeuLysThr-AspGlyHisGlnGlnPhePheAlaIrl	224
Db	2	TTCAATGTTAGTCTTAGAGAAACAGGAAAAATACTGATGGTCACCGACGAGTTCTTCGCGAAT	61
Qy	224	eValGlnLeuIleGlyThrArgLysGlnAlaGluAsnPheAlaTrArgLeuGluLeuAs	244
Db	62	CGTACAGCTGATAGAAACACCGACGACGCTGAATAATTTTGTCTTACCGACTTGACGTAAA	121
Qy	244	nGlyHisArgArgLeuThrTrpGluAlaThrProArgSerIleHisGluGlyIleAl	264
Db	122	TGGTCATAGCGCAGATTGACTTTGGGAAGCGACTCCTCGATCTATTATGAAGGAATGCG	181
Qy	264	aThrAlaIleMetAsnSerAspCysLeuValPheAspThrSerIleAlaGlnLeuPheAl	284
Db	182	AACAGCCATTATGAATACGACTGTCTAGTCTTTTGACACCGACATTGCGACGCTTTTTCG	241
Qy	284	aGluAsnGlyAsnLeuGlyIleAsnValThrIleSerMetCys	298
Db	242	AGAAAATGGCAATTTAGGCATCAATGTAACTATTTCATTTCCATGTGT	284

RESULT 15

```

US-10-437-963-83416/C
; Sequence 83416, Application US/10437963
; Publication NO. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules a
; TITLE OF INVENTION: Plants and Uses Thereof for
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 83416
; LENGTH: 1438
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_8274C.1
US-10-437-963-83416

```

```

Alignment Scores:
Pred. No.: 11e-46 Length: 1438
Score: 462.50 Matches: 111
Percent Similarity: 54.96% Conservative: 44
Best Local Similarity: 39.36% Mismatches: 101
Query Match: 28.71% Indels: 27
DB: 18 Gaps: 10

US-10-679-246-2 (1-298) x US-10-437-963-83416 (1-1438)

QY 28 GlyThrSerIysCysProPro-----SerGlnArgValProAlaLeu 41
||| ||| ||| ||| : : :
Db 1154 GGAGATTTCATGTCGAAGCCCTGGTCTACATCACTTGTCTACAGTTGAGCTATCATCGCTG 1095

QY 42 ThrGlyThrThrAlaSerAsnAsnAspLeuAlaSerLeuPheGluCysProValCysPhe 61
||| ||| ||| ||| |||
Db 1094 ACGGGC-----TTGATGATTTGCTCGAGTGTCCAGTCTGCACA 1056

QY 62 AspTyrValLeuProProlLeuGlnCysGlnSerGlyHisLeuValCysSerAsnCys 81
||| : : : ||| : : :
Db 1055 AACTCAATGGCCCACTATCTACTGAGTGGCCAAATGGTCACAAATATGCTCTAACTGC 996

```

```
QY 82 ArgProLysLeu---ThrCysCysProThrCysArgGlyProLeuGlySerIleArgAsn 100
Db   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 995 AAGCATAGGTGGAGAAATCATTTGCCCTACTTGTGCCAAGAACTGGGTAATATCAGATGT 936
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 101 LeuAlaMetGluLysValAlaAsnSerValLeuPheProCysLysTyrAlaSerSerGly 120
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 935 TTGGCTCTTGAAGGTGGCTGAGTCACCTTCCATGCAAAATATCAAAAGCCTAGGC 876
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 CysGluIleThrLeuProHisThrGluLysAlaAspHisGluGluLeuCysGluPheArg 140
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 875 TGTGCTGAGATTCATCCCTTACCAAAACAACTTAAGCATGAGGAGCTTCGAGGTTCCAG 816
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 141 ProTyrSerCysProCysProGlyAlaSerCysLysTyrPngInGlySerLeuAspAlaVal 160
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 815 CCATATAGTTGTCATATGACAGTTCCAGATGCCGATGTTGCTGGCGATGTTCCGATGCTT 756
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 161 MetProHisLeuMetHisGlnHisLysSerIleThrThrLeuGlnGlyGluAspIleVal 180
Db   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 755 GTCTCTCATCTCATAAACGACCATAAAG--GTGGACTTGCATGAAGGCTGCACCTTTAAC 699
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 PheLeuAlaThrAspIleAsnLeuProGlyAlaValAsp-----TrpValMet--- 196
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 698 CACCGTTATGTGAAGTCCAAAC--CCTATGAAGTGGAAACGCTACATGGATGCTGACG 642
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 197 MetGlnSerCysPheGlyPheHisPheMetLeu-ValLeuGluLysGlnGluLysTyrAs 216
Db   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 641 GTTTTCAAGTGTITGGACAGCACTTCTGCCTGCATTGTG---AGGCATTCTCTGTGGGG 585
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 216 pGlyHisGlnGlnPhePheAlaIleValGlnLeuIleGlyThrArgLysGlnAlaGluAs 236
Db   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 584 ATGGCACCAGTG-TACATGGCGTCTCTGAGGTTTCATGGCGGAGGACAGCGAGGCCCGGAA 526
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 236 nPheAlaTyrArgLeuGluLeuAsnGlyHisArgArgArgLeuThrTrpGluAlaThrPr 256
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 525 CTTCGTACACGCTGAGGTGGCGGGAACGGCGGAAGCTGACATGGCAGGGCATCCC 466
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 256 oArgSerIleHisGluGlyIleAlaThrAlaIleMetAsnSerAspCysLeuValPheAs 276
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 465 CCGGAGCATCAGGACAGCCACAGAGAGTCCGGGACAGCTTCGACGGCTCATCATCCA 406
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 276 pThrSerIleAlaGlnLeuPheAlaGluAsnGlyAsn-----LeuGlyIleAsnVa 293
Db   ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 405 CCGGAACATGGCC--CTCTTCTCTCCGGCGGCAACAGGCGAGGAGCTCAAGCTGGGGT 349
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 293 lThr 294
Db   |||||
QY 348 TACC 345
Db   |||||
```

Search completed: April 25, 2005, 10:57:04  
Job time : 673 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 22:57:31 ; Search time 5658 Seconds  
(without alignments)  
10910.559 Million cell updates/sec

Title: US-10-679-246-1  
Perfect score: 1274  
Sequence: 1 ttcttttgtagtttatgtg.....ctctttcgtagtggaagc 1274

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 15

Total number of hits satisfying chosen parameters: 122303

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sta.\*

12: gb.sv.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1274	100.0	1274	6	AR411820 Sequence
2	1274	100.0	1274	6	AX058082 Sequence
3	1274	100.0	2924	6	CQ491092 Sequence
4	1274	100.0	2924	6	CQ493590 Sequence
5	1274	100.0	2924	6	CQ496955 Sequence
6	1274	100.0	31705	9	HSB400626 Homo sapi
7	1274	100.0	173304	9	AC023818 Homo sapi
8	1223	96.0	2829	6	AX833145 Sequence
9	1223	96.0	2829	9	AK094663 Homo sapi
10	1172	92.0	2972	9	HSB807215 Homo sapi
11	1104	86.7	2128	9	AK056051 Homo sapi
12	1041	81.7	2440	6	AR380043 Sequence
13	958	75.2	1540	6	BC035562 Homo sapi
14	958	75.2	1874	6	CQ719559 Sequence
15	958	75.2	2048	9	HSU76247 Human hSIAB
16	957	75.1	2454	9	BC042550 Homo sapi
17	947	74.3	1886	9	HSU63295 Homo sapien
18	881	69.2	4090	6	AX780380 Sequence
19	838	65.8	2034	9	BC018193 Homo sapi

20	780	61.2	1884	6	A63558	A63558 Sequence 11
21	780	61.2	1884	6	AR271267	AR271267 Sequence
c 22	159	12.5	405	6	CQ506573	CQ506573 Sequence
c 23	154	12.1	277	6	CQ476643	CQ476643 Sequence
c 24	142	11.1	673	11	BV208831	BV208831 SIFAHI 234
c 25	139	10.9	417	6	CQ511941	CQ511941 Sequence
c 26	128	10.0	473	6	CQ516613	CQ516613 Sequence
c 27	118	9.3	411	6	CQ481952	CQ481952 Sequence
c 28	118	9.3	441	6	CQ503091	CQ503091 Sequence
c 29	92	7.2	523	11	BV102520	BV102520 MARC 1442
c 30	88	6.9	194622	10	AC133654	AC133654 Mus muscu
c 31	85	6.7	438	6	CQ472783	CQ472783 Sequence
c 32	77	6.0	1845	10	BC046317	BC046317 Mus muscu
c 33	77	6.0	1968	6	AX305603	AX305603 Sequence
c 34	77	6.0	1968	10	MMSIAH1A	Z19579 M.musculus
c 35	77	6.0	175636	10	AC142211	AC142211 Mus muscu
c 36	62	4.9	732	11	BV104504	BV104504 MARC 1442
c 37	62	4.9	732	11	BV104504	BV104504 MARC 1442
c 38	53	4.2	1457	10	AB067814	AB067814 Rattus no
c 39	53	4.2	1465	10	AF389476	AF389476 Rattus no
c 40	53	4.2	220157	2	AC105462	AC105462 Rattus no
c 41	53	4.2	232382	2	AC098991	AC098991 Rattus no
c 42	52	4.1	426	6	CQ486802	CQ486802 Sequence
c 43	52	4.1	2123	10	MMSIAHPSA	Z19582 M.musculus
c 44	52	4.1	168091	10	AC139553	AC139553 Mus muscu
c 45	52	4.1	182733	10	AC123941	AC123941 Mus muscu
c 46	45	3.5	1713	10	MMSIAH1B	Z19580 M.musculus
c 47	45	3.5	1720	10	BC052887	BC052887 Mus muscu
c 48	45	3.5	112893	10	AL732294	AL732294 Mouse DNA
c 49	45	3.5	234105	10	AC091606	AC091606 Mus Muscu
c 50	37	2.9	523	11	BV102520	BV102520 MARC 1442
c 51	35	2.7	196	10	MMSIAHPBS	Z19583 M.musculus
c 52	35	2.7	600	5	AF411976	AF411976 Gallus ga
c 53	35	2.7	711	5	CR386772	CR386772 Gallus ga
c 54	32	2.5	89948	9	AL359734	AL359734 Human DNA
c 55	30	2.4	1212	10	AF389477	AF389477 Rattus no
c 56	30	2.4	2048	10	AB067815	AB067815 Rattus no
c 57	30	2.4	2198	6	AX305605	AX305605 Sequence
c 58	30	2.4	2198	10	MMSIAH2A	Z19581 M.musculus
c 59	30	2.4	2511	10	BC058400	BC058400 Mus muscu
c 60	30	2.4	164715	2	AC142228	AC142228 Mus muscu
c 61	30	2.4	181603	2	AC119873	AC119873 Mus muscu
c 62	30	2.4	221589	2	AC126910	AC126910 Rattus no
c 63	30	2.4	254091	2	AC112531	AC112531 Rattus no
c 64	28	2.2	1509	5	AF155509	AF155509 Xenopus l
c 65	26	2.0	1246	5	BC072747	BC072747 Xenopus l
c 66	25	2.0	280815	2	AC107562	AC107562 Rattus no
c 67	24	1.9	9462	3	AF466306	AF466306 Dictyoste
c 68	24	1.9	129435	9	AL391314	AL391314 Human DNA
c 69	24	1.9	140179	2	AC084688	AC084688 Homo sapi
c 70	24	1.9	211456	10	AL928926	AL928926 Mouse DNA
c 71	23	1.8	341	5	AF529483	AF529483 Acipenser
c 72	23	1.8	41352	3	AV190936	AV190936 Drosophil
c 73	22	1.7	143119	8	AP005068	AP005068 Oryza sat
c 74	22	1.7	149628	2	AC151175	AC151175 Bos tauru
c 75	22	1.7	160788	9	AC128653	AC128653 Homo sapi
c 76	22	1.7	181870	2	AC068474	AC068474 Homo sapi
c 77	22	1.7	184121	9	AC023119	AC023119 Homo sapi
c 78	22	1.7	186272	2	AC134495	AC134495 Rattus no
c 79	22	1.7	195419	2	CR550302	CR550302 Danio rer
c 80	22	1.7	204618	2	EX321889	EX321889 Danio rer
c 81	22	1.7	217636	10	AC127411	AC127411 Mus muscu
c 82	22	1.7	222312	2	AC122624	AC122624 Rattus no
c 83	22	1.7	227154	10	AC115121	AC115121 Mus muscu
c 84	22	1.7	235733	2	AC106933	AC106933 Rattus no
c 85	21	1.6	810	5	CR390245	CR390245 Gallus ga
c 86	21	1.6	1094	6	CQ771162	CQ771162 Sequence
c 87	21	1.6	1150	6	CQ771161	CQ771161 Sequence
c 88	21	1.6	1913	9	AK122746	AK122746 Homo sapi
c 89	21	1.6	5047	9	AF283519S7	AF283519S7 Homo sapi
c 90	21	1.6	35056	3	CER10H10	Z70586 Caenorhabdi
c 91	21	1.6	72735	8	AP004487	AP004487 Lotus cor
c 92	21	1.6	94430	8	AC149544	AC149544 Populus b

93	21	1.6	110000	2	AC109084.1	Continuation (2 of	166	20	1.6	52454	2	AC100936	AC100936 Mus muscu
94	21	1.6	116305	9	AC106876	Homo sapi	c 167	20	1.6	55339	9	AC093593	AC093593 Homo sapi
95	21	1.6	121764	8	AC149804	Medicago	168	20	1.6	56031	9	AL157377	AL157377 Human DNA
96	21	1.6	139939	8	AC008965	Homo sapi	169	20	1.6	77835	2	PFMAL13P2.3	Continuation (4 of
97	21	1.6	140031	2	AC121686	AC121686 Rattus no	c 170	20	1.6	78794	9	AC003990	AC003990 Homo sapi
98	21	1.6	149724	10	AC136637	AC136637 Mus muscu	171	20	1.6	84313	3	AC103565	AC103565 Caenorhab
99	21	1.6	157217	2	BNX36324	Danio rer	c 172	20	1.6	84367	3	AC005454	AC005454 Drosophil
100	21	1.6	160673	8	CNS080DR	AC133185 Mus muscu	c 173	20	1.6	84550	3	PFMAL1P2.3	Continuation (4 of
101	21	1.6	163888	10	AC133185	AC133185 Mus muscu	174	20	1.6	84710	5	AL928945	AL928945 Zebrafish
102	21	1.6	165200	5	BSX323079	BSX323079 Zebrafish	175	20	1.6	88412	2	AC141373	AC141373 Rattus no
103	21	1.6	165675	9	AC022163	AC022163 Homo sapi	c 176	20	1.6	89652	5	BSX24196	BSX24196 Zebrafish
104	21	1.6	166001	9	BSA462D18	BSA462D18 Human DNA	c 177	20	1.6	89755	9	AC173585	AC173585 Homo sapi
105	21	1.6	173926	5	BSX005042	BSX005042 Zebrafish	178	20	1.6	91587	2	AC112924	AC112924 Mus muscu
106	21	1.6	180483	9	AC122707	AC122707 Homo sapi	c 179	20	1.6	93853	2	CR3881612	Continuation (4 of
107	21	1.6	183483	9	AL161628	AL161628 Human DNA	c 180	20	1.6	100241	2	CR3881612	CR3881612 Danio rer
108	21	1.6	190869	9	AC120393	AC120393 Mus muscu	181	20	1.6	102306	5	BSX71980	BSX71980 Zebrafish
109	21	1.6	196203	9	AL355355	AL355355 Human DNA	c 182	20	1.6	110000	1	AE017332.7	AE017332.7
110	21	1.6	205476	9	AL355355	AL355355 Sus scrof	c 183	20	1.6	110000	1	CR000013.3	Continuation (8 of
111	21	1.6	208202	2	AC013446	AC013446 Homo sapi	c 184	20	1.6	110000	2	AC106120.2	Continuation (4 of
112	21	1.6	215147	10	AC119806	AC119806 Mus muscu	c 185	20	1.6	110000	2	AC111970.4	Continuation (3 of
113	21	1.6	215325	2	AC114171	AC114171 Rattus no	c 186	20	1.6	110000	2	AC126295.1	Continuation (5 of
114	21	1.6	222866	2	AC131557	AC131557 Rattus no	c 187	20	1.6	110000	2	BSX649565	BSX649565 Mus muscu
115	21	1.6	224117	2	AC129992	AC129992 Rattus no	c 188	20	1.6	110000	8	CR382127.03	Continuation (4 of
116	21	1.6	226909	2	AC079490	AC079490 Mus muscu	189	20	1.6	119526	8	CR382127.22	Continuation (23 o
117	21	1.6	228597	2	AC107556	AC107556 Rattus no	c 190	20	1.6	121506	4	AC087807	AC087807 Felis cat
118	21	1.6	233345	4	AJ421481	AJ421481 Bos tauru	c 191	20	1.6	124070	2	AC136159	AC136159 Rattus n
119	21	1.6	235926	2	AC136859	AC136859 Rattus no	c 192	20	1.6	124595	2	AC151399	AC151399 Dasytus n
120	21	1.6	241073	2	AC095479	AC095479 Rattus no	193	20	1.6	125681	10	AC122889	AC122889 Mus muscu
121	21	1.6	243500	2	AC100177	AC100177 Mus muscu	194	20	1.6	125681	10	AC122889	AC122889 Mus muscu
122	21	1.6	249703	2	AC126988	AC126988 Rattus no	195	20	1.6	125681	10	AC122889	AC122889 Mus muscu
123	21	1.6	251870	2	AC118369	AC118369 Rattus no	c 196	20	1.6	125681	10	AC122889	AC122889 Mus muscu
124	21	1.6	259384	2	AC120595	AC120595 Rattus no	c 197	20	1.6	125681	10	AC122889	AC122889 Mus muscu
125	21	1.6	260224	6	AR411827	AR411827 Sequence	198	20	1.6	125681	10	AC122889	AC122889 Mus muscu
126	21	1.6	260224	6	AR411827	AR411827 Sequence	199	20	1.6	125681	10	AC122889	AC122889 Mus muscu
127	21	1.6	260224	6	AR411827	AR411827 Sequence	200	20	1.6	125681	10	AC122889	AC122889 Mus muscu
128	21	1.6	260224	6	AR411827	AR411827 Sequence	201	20	1.6	125681	10	AC122889	AC122889 Mus muscu
129	21	1.6	260224	6	AR411827	AR411827 Sequence	202	20	1.6	125681	10	AC122889	AC122889 Mus muscu
130	21	1.6	260224	6	AR411827	AR411827 Sequence	203	20	1.6	125681	10	AC122889	AC122889 Mus muscu
131	21	1.6	260224	6	AR411827	AR411827 Sequence	204	20	1.6	125681	10	AC122889	AC122889 Mus muscu
132	21	1.6	260224	6	AR411827	AR411827 Sequence	205	20	1.6	125681	10	AC122889	AC122889 Mus muscu
133	21	1.6	260224	6	AR411827	AR411827 Sequence	206	20	1.6	125681	10	AC122889	AC122889 Mus muscu
134	21	1.6	260224	6	AR411827	AR411827 Sequence	207	20	1.6	125681	10	AC122889	AC122889 Mus muscu
135	21	1.6	260224	6	AR411827	AR411827 Sequence	208	20	1.6	125681	10	AC122889	AC122889 Mus muscu
136	21	1.6	260224	6	AR411827	AR411827 Sequence	209	20	1.6	125681	10	AC122889	AC122889 Mus muscu
137	21	1.6	260224	6	AR411827	AR411827 Sequence	210	20	1.6	125681	10	AC122889	AC122889 Mus muscu
138	21	1.6	260224	6	AR411827	AR411827 Sequence	211	20	1.6	125681	10	AC122889	AC122889 Mus muscu
139	21	1.6	260224	6	AR411827	AR411827 Sequence	212	20	1.6	125681	10	AC122889	AC122889 Mus muscu
140	21	1.6	260224	6	AR411827	AR411827 Sequence	213	20	1.6	125681	10	AC122889	AC122889 Mus muscu
141	21	1.6	260224	6	AR411827	AR411827 Sequence	214	20	1.6	125681	10	AC122889	AC122889 Mus muscu
142	21	1.6	260224	6	AR411827	AR411827 Sequence	215	20	1.6	125681	10	AC122889	AC122889 Mus muscu
143	21	1.6	260224	6	AR411827	AR411827 Sequence	216	20	1.6	125681	10	AC122889	AC122889 Mus muscu
144	21	1.6	260224	6	AR411827	AR411827 Sequence	217	20	1.6	125681	10	AC122889	AC122889 Mus muscu
145	21	1.6	260224	6	AR411827	AR411827 Sequence	218	20	1.6	125681	10	AC122889	AC122889 Mus muscu
146	21	1.6	260224	6	AR411827	AR411827 Sequence	219	20	1.6	125681	10	AC122889	AC122889 Mus muscu
147	21	1.6	260224	6	AR411827	AR411827 Sequence	220	20	1.6	125681	10	AC122889	AC122889 Mus muscu
148	21	1.6	260224	6	AR411827	AR411827 Sequence	221	20	1.6	125681	10	AC122889	AC122889 Mus muscu
149	21	1.6	260224	6	AR411827	AR411827 Sequence	222	20	1.6	125681	10	AC122889	AC122889 Mus muscu
150	21	1.6	260224	6	AR411827	AR411827 Sequence	223	20	1.6	125681	10	AC122889	AC122889 Mus muscu
151	21	1.6	260224	6	AR411827	AR411827 Sequence	224	20	1.6	125681	10	AC122889	AC122889 Mus muscu
152	21	1.6	260224	6	AR411827	AR411827 Sequence	225	20	1.6	125681	10	AC122889	AC122889 Mus muscu
153	21	1.6	260224	6	AR411827	AR411827 Sequence	226	20	1.6	125681	10	AC122889	AC122889 Mus muscu
154	21	1.6	260224	6	AR411827	AR411827 Sequence	227	20	1.6	125681	10	AC122889	AC122889 Mus muscu
155	21	1.6	260224	6	AR411827	AR411827 Sequence	228	20	1.6	125681	10	AC122889	AC122889 Mus muscu
156	21	1.6	260224	6	AR411827	AR411827 Sequence	229	20	1.6	125681	10	AC122889	AC122889 Mus muscu
157	21	1.6	260224	6	AR411827	AR411827 Sequence	230	20	1.6	125681	10	AC122889	AC122889 Mus muscu
158	21	1.6	260224	6	AR411827	AR411827 Sequence	231	20	1.6	125681	10	AC122889	AC122889 Mus muscu
159	21	1.6	260224	6	AR411827	AR411827 Sequence	232	20	1.6	125681	10	AC122889	AC122889 Mus muscu
160	21	1.6	260224	6	AR411827	AR411827 Sequence	233	20	1.6	125681	10	AC122889	AC122889 Mus muscu
161	21	1.6	260224	6	AR411827	AR411827 Sequence	234	20	1.6	125681	10	AC122889	AC122889 Mus muscu
162	21	1.6	260224	6	AR411827	AR411827 Sequence	235	20	1.6	125681	10	AC122889	AC122889 Mus muscu
163	21	1.6	260224	6	AR411827	AR411827 Sequence	236	20	1.6	125681	10	AC122889	AC122889 Mus muscu
164	21	1.6	260224	6	AR411827	AR411827 Sequence	237	20	1.6	125681	10	AC122889	AC122889 Mus muscu
165	21	1.6	260224	6	AR411827	AR411827 Sequence	238	20	1.6	125681	10	AC122889	AC122889 Mus muscu

C 239	20	1.6 172336	9	CNS06C7Z	AL391153 Human chr	C 312	20	1.6 244180	2	AC094031	AC094031 Rattus no
240	20	1.6 173633	2	AC107793	Mus muscu	C 313	20	1.6 244598	2	AC105871	Rattus no
241	20	1.6 173740	2	AC020964	Mus muscu	C 314	20	1.6 245413	2	AC095423	Rattus no
C 242	20	1.6 173769	9	AC113933	Homo sapi	315	20	1.6 246742	2	AC103033	Rattus no
243	20	1.6 174037	9	AC009995	Homo sapi	316	20	1.6 247845	2	AC094039	Rattus no
244	20	1.6 174540	8	AC068801	Homo sapi	317	20	1.6 248001	2	AC051614	Mus muscu
C 245	20	1.6 176349	8	AP002092	Oryza sat	318	20	1.6 248178	2	AC097387	Rattus no
246	20	1.6 177488	8	AC146766	Oryza sat	C 319	20	1.6 249107	2	AC094049	Rattus no
247	20	1.6 177613	10	AC101968	Mus muscu	C 320	20	1.6 252240	3	AE014841	Plasmodiu
248	20	1.6 178142	2	AC133840	Rattus no	C 321	20	1.6 252524	2	AC114081	Rattus no
249	20	1.6 178733	9	AC026003	Homo sapi	322	20	1.6 252749	2	AC096611	Rattus no
250	20	1.6 179941	2	AC145934	Gallus ga	323	20	1.6 253780	2	AC113716	Rattus no
251	20	1.6 180136	1	BAC180K	D26185 B. subtilis	324	20	1.6 265677	5	CNS09SBK	Clone A24
252	20	1.6 180714	8	AP005804	Oryza sat	325	20	1.6 265677	5	CNS09SBP	AE052609 Tetraodon
C 253	20	1.6 181096	10	AL592187	Mouse DNA	326	20	1.6 269102	3	AE003628	AE003628 Drosophila
254	20	1.6 182371	10	AC147263	Mus muscu	C 327	20	1.6 269797	2	AC120567	Rattus no
255	20	1.6 183156	2	AC023062	Mus muscu	328	20	1.6 269854	2	AC107448	Rattus no
C 256	20	1.6 183958	2	BX530096	BX530096 Mus muscu	329	20	1.6 270153	2	AC096262	Rattus no
257	20	1.6 187984	2	AP001262	AP001262 Homo sapi	330	20	1.6 274417	2	AC128979	Rattus no
C 258	20	1.6 188008	10	AC078863	Mus muscu	C 331	20	1.6 275351	2	AC117883	Rattus no
259	20	1.6 188218	10	AL136998	Mouse DNA	C 332	20	1.6 280240	2	AC106683	Rattus no
C 260	20	1.6 188937	10	AC108433	Mus muscu	C 333	20	1.6 281089	2	AC094473	Rattus no
261	20	1.6 189736	6	CO870483	Sequence	C 334	20	1.6 284886	2	AC110112	Rattus no
262	20	1.6 190771	2	AC140346	Mus muscu	335	20	1.6 289591	2	AC125748	Rattus no
263	20	1.6 191219	2	AC109499	Mus muscu	C 336	20	1.6 297657	2	AC120626	Rattus no
C 264	20	1.6 191396	2	AC118776	AC118776 Rattus no	337	20	1.6 301278	1	AE015939	AE015939 Clostridi
265	20	1.6 191735	10	AC092711	AP001267 Homo sapi	C 338	20	1.6 302000	1	AP003187	AP003187 Leptospir
C 266	20	1.6 194310	9	AP001267	Mus muscu	C 339	20	1.6 304923	1	AE017289	Rattus no
C 267	20	1.6 194449	2	AC103673	Mus muscu	C 340	20	1.6 309014	2	AC129655	Rattus no
268	20	1.6 194602	2	AC124087	AC124087 Mus muscu	C 341	20	1.6 310003	1	AE014016	AE014016 Buchnera
C 269	20	1.6 194849	2	AC147015	AC147015 Zea mays	C 342	20	1.6 312241	2	AC108636	Rattus no
C 270	20	1.6 195305	10	AL713967	Mouse DNA	343	20	1.6 323222	2	AC112459	Rattus no
C 271	20	1.6 195814	2	AC129635	AC129635 Rattus no	344	20	1.6 328550	2	AC142772	Macaca mu
C 272	20	1.6 196803	10	AC132461	AC132461 Mus muscu	C 345	20	1.6 329753	3	AC116986	Dictyoste
C 273	20	1.6 197943	2	AC129321	AC129321 Mus muscu	346	20	1.6 333321	2	AC134498	Rattus no
C 274	20	1.6 198057	2	AC126072	AC126072 Rattus no	347	20	1.6 346597	2	AC134498	Rattus no
275	20	1.6 198573	10	AL929186	AL929186 Mouse DNA	C 348	20	1.6 346597	2	AC134498	Rattus no
C 276	20	1.6 201137	2	AC128312	AC128312 Rattus no	C 349	20	1.6 349980	6	Q0870482	Q0870482 Sequence
C 277	20	1.6 203207	2	AC141523	AC141523 Rattus no	350	20	1.6 349980	6	AX344559	AX344559 Sequence
C 278	20	1.6 206503	2	BX324197	BX324197 Danio rer	C 351	20	1.6 349980	6	AX344560	AX344560 Sequence
C 279	20	1.6 207008	2	CR847513	CR847513 Danio rer	C 352	19	1.5 19	6	AR411828	Sequence
280	20	1.6 207107	2	CR354437	CR354437 Danio rer	C 353	19	1.5 19	6	AX058124	Sequence
281	20	1.6 208140	2	CR255953	CR255953 Danio rer	C 354	19	1.5 231	6	Q0469285	Q0469285 Sequence
282	20	1.6 208770	2	CR388155	CR388155 Danio rer	355	19	1.5 364	8	ATH530949	ATH530949 Arabidops
C 283	20	1.6 209326	10	AC110247	AC110247 Mus muscu	356	19	1.5 414	9	HUM2C18X04	HUM2C18X04
C 284	20	1.6 210427	10	AC124757	AC124757 Mus muscu	C 357	19	1.5 447	6	Q0499651	Q0499651 Sequence
C 285	20	1.6 211167	2	AC091249	AC091249 Mus muscu	C 358	19	1.5 447	6	Q0508621	Q0508621 Sequence
C 286	20	1.6 211381	10	AC116404	AC116404 Mus muscu	C 359	19	1.5 470	6	Q0515194	Q0515194 Sequence
C 287	20	1.6 212159	2	AC101877	AC101877 Mus muscu	C 360	19	1.5 471	14	AY573812	AY573812 Pneumonia
288	20	1.6 213080	1	BSUB0001	Z99104 Bacillus su	C 361	19	1.5 601	11	BV191886	BV191886 sgmm17248
C 289	20	1.6 213862	10	AC126552	AC126552 Mus muscu	C 362	19	1.5 620	11	G85128	G85128 S208P6441RE
290	20	1.6 214250	9	AC108050	AC108050 Homo sapi	C 363	19	1.5 702	11	BV039848	BV039848 S212P6344
291	20	1.6 214290	5	BX294379	BX294379 Zebrafish	C 364	19	1.5 850	6	Q0478453	Q0478453 Sequence
292	20	1.6 214308	10	AL590870	AL590870 Mouse DNA	365	19	1.5 939	10	AY318104	Mus muscu
293	20	1.6 215638	10	AC124200	AC124200 Mus muscu	366	19	1.5 948	6	Q0728097	Q0728097 Sequence
294	20	1.6 217046	10	AC122389	AC122389 Mus muscu	C 367	19	1.5 978	10	AY073718	AY073718 Mus muscu
295	20	1.6 218573	10	AC111460	AC111460 Rattus no	368	19	1.5 1015	6	AX0702766	AX0702766 Sequence
C 296	20	1.6 219812	10	AC118196	AC118196 Mus muscu	C 369	19	1.5 1065	1	AF448128	AF448128 Clostridi
297	20	1.6 220894	10	AL732410	AL732410 Mouse DNA	C 370	19	1.5 1065	1	AF448129	AF448129 Clostridi
C 298	20	1.6 221557	2	AC146538	AC146538 Gasterost	C 371	19	1.5 1069	3	AB057568	AB057568 Carogloss
C 299	20	1.6 223109	2	BX957356	BX957356 Danio rer	372	19	1.5 1375	5	AF003840	AF003840 Hydrolagu
300	20	1.6 224643	2	AC097155	AC097155 Rattus no	373	19	1.5 1473	6	AR399333	AR399333 Sequence
301	20	1.6 225152	10	AC132454	AC132454 Mus muscu	374	19	1.5 1473	6	AR399333	AR399333 Sequence
C 302	20	1.6 228961	2	AC103135	AC103135 Rattus no	375	19	1.5 1800	9	BC069666	BC069666 Homo sapi
303	20	1.6 229791	2	CR391941	CR391941 Danio rer	C 376	19	1.5 1807	3	PFU07365	PFU07365 Plasmodium
304	20	1.6 230701	10	AL713871	AL713871 Mouse DNA	C 377	19	1.5 1884	1	AF282878	AF282878 Pseudomon
305	20	1.6 231276	2	AC128991	AC128991 Rattus no	C 378	19	1.5 1946	3	DR081164	DR081164 Drosophila
306	20	1.6 232617	2	AC135871	AC135871 Rattus no	C 379	19	1.5 1946	3	DR081164	DR081164 Drosophila
307	20	1.6 237164	2	AC114255	AC114255 Rattus no	C 380	19	1.5 1951	6	Q0508057	Q0508057 Sequence
C 308	20	1.6 237165	10	AC112948	AC112948 Mus muscu	381	19	1.5 1995	6	AR380510	AR380510 Sequence
309	20	1.6 237378	2	AC087038	AC087038 Mus muscu	382	19	1.5 1995	6	HUMCYP2C18	HUMCYP2C18
310	20	1.6 237754	10	AC125576	AC125576 Rattus no	383	19	1.5 2000	6	AX655140	AX655140 Sequence
C 311	20	1.6 241249	10	AL935147	AL935147 Mouse DNA	384	19	1.5 2009	6	AR071577	AR071577 Sequence

C 385	19	1.5	2136	10	RNU37539	U37539 Rattus norv	458	19	1.5	40874	9	AF129108	AF129108 Homo sapi
C 386	19	1.5	2136	10	RNU37540	U37540 Rattus norv	C 459	19	1.5	41516	9	AX470067	AX470067 Human DNA
C 387	19	1.5	2258	6	AR071580	Sequence	C 460	19	1.5	41836	9	AC021052	AC021052 Homo sapi
C 388	19	1.5	2258	6	AR380450	Sequence	C 461	19	1.5	43379	3	CEC02C6	Z79596 Caenorhabdi
C 389	19	1.5	2258	6	AX409722	Sequence	C 462	19	1.5	43944	3	AY642135	AY642135 Homo sapi
C 390	19	1.5	2258	6	HUM2C18	Sequence	C 463	19	1.5	44521	3	AF045640	AF045640 Caenorhab
C 391	19	1.5	2271	1	CDI300676	Sequence	C 464	19	1.5	46650	9	AL358783	AL358783 Human DNA
C 392	19	1.5	2271	6	AX259458	Sequence	C 465	19	1.5	47880	3	AC006651	AC006651 Caenorhab
C 393	19	1.5	2271	6	AX512904	Sequence	C 466	19	1.5	51340	2	AC025213	AC025213 Homo sapi
C 394	19	1.5	2306	14	PNV1CBNP	D10331 Pneumonia v	C 467	19	1.5	55163	2	AC020517	AC020517 Drosophill
C 395	19	1.5	2395	9	HUMSVF450A	L07093 Human cytoc	C 468	19	1.5	57969	2	AC100083	AC100083 Mus muscu
C 396	19	1.5	2396	8	AY054495	Arabidops	C 469	19	1.5	58149	2	AC102560	AC102560 Mus muscu
C 397	19	1.5	2396	8	DMU34039	U34039 Drosophilla	C 470	19	1.5	58822	6	AX695431	AX695431 Sequence
C 398	19	1.5	2520	8	AY369264	AY369264 Cochliob	C 471	19	1.5	59832	1	CR628339	CR628339 Legionell
C 399	19	1.5	2520	3	DDGLPV	X62142 D. discoide	C 472	19	1.5	60945	9	AC079394	AC079394 Homo sapi
C 400	19	1.5	2538	10	RNNAC	X78034 R. norvegicu	C 473	19	1.5	62190	9	AC116301	AC116301 Homo sapi
C 401	19	1.5	2599	10	AF112187	AF112187 Mus muscu	C 474	19	1.5	63998	2	AC101027	AC101027 Mus muscu
C 402	19	1.5	3001	6	BD265989	BD265989 Schizophr	C 475	19	1.5	65199	2	AC100669	AC100669 Mus muscu
C 403	19	1.5	3001	6	AR244519	AR244519 Sequence	C 476	19	1.5	66066	8	AC149486	AC149486 Populus b
C 404	19	1.5	3012	10	RNENACG	X77933 R. norvegicu	C 477	19	1.5	66262	9	AC132516	AC132516 Homo sapi
C 405	19	1.5	3018	10	BC021338	BC021338 Mus muscu	C 478	19	1.5	66367	2	AX119990	AX119990 Continuation (4 of
C 406	19	1.5	3227	5	CQ583949	CQ583949 Sequence	C 479	19	1.5	68256	8	AC149543	AC149543 Populus b
C 407	19	1.5	3236	5	AY394964	AY394964 Danio rer	C 480	19	1.5	69830	2	AC100204	AC100204 Mus muscu
C 408	19	1.5	3545	5	AF153446	AF153446 Danio rer	C 481	19	1.5	72334	2	AP000677	AP000677 Homo sapi
C 409	19	1.5	3658	5	BC054623	BC054623 Danio rer	C 482	19	1.5	73834	2	AC020405	AC020405 Drosophill
C 410	19	1.5	3662	6	CQ872871	CQ872871 Sequence	C 483	19	1.5	74573	9	AL589988	AL589988 Human DNA
C 411	19	1.5	3888	10	D85517	D85517 Mesocricetu	C 484	19	1.5	75013	10	AX088537	AX088537 Mouse DNA
C 412	19	1.5	3991	6	CQ581372	CQ581372 Sequence	C 485	19	1.5	75073	3	CEY48A6B	CEY48A6B Caenorhab
C 413	19	1.5	4026	6	CQ580856	CQ580856 Sequence	C 486	19	1.5	75519	2	AC124999	AC124999 Mus muscu
C 414	19	1.5	4259	3	AB091123	AB091123 Haemocynth	C 487	19	1.5	75519	2	AC124999	AC124999 Mus muscu
C 415	19	1.5	4340	8	AF213504	AF213504 Brassica	C 488	19	1.5	76388	2	CR450853	CR450853 Continuation (4 of
C 416	19	1.5	4372	6	CQ581414	CQ581414 Sequence	C 489	19	1.5	76388	2	AC100200	AC100200 Mus muscu
C 417	19	1.5	4596	3	AF310896	AF310896 Dictyoste	C 490	19	1.5	78000	5	AL645795	AL645795 Zebrafish
C 418	19	1.5	4711	3	AF238314	AF238314 Dictyoste	C 491	19	1.5	80288	9	AC004901	AC004901 Homo sapi
C 419	19	1.5	4821	6	CQ872839	CQ872839 Sequence	C 492	19	1.5	81236	8	AC007168	AC007168 Arabidops
C 420	19	1.5	5107	9	AF250332	AF250332 Homo sapi	C 493	19	1.5	82659	9	AC004553	AC004553 Homo sapi
C 421	19	1.5	6131	6	AX458508	AX458508 Sequence	C 494	19	1.5	83351	8	AC010657	AC010657 Human DNA
C 422	19	1.5	6641	6	AX344595	AX344595 Sequence	C 495	19	1.5	83351	8	AC010657	AC010657 Genomic s
C 423	19	1.5	7597	6	AX345915	AX345915 Sequence	C 496	19	1.5	83698	9	AL731683	AL731683 Human DNA
C 424	19	1.5	7728	6	AX344979	AX344979 Sequence	C 497	19	1.5	83632	9	AL445470	AL445470 Human DNA
C 425	19	1.5	7728	6	AX348500	AX348500 Sequence	C 498	19	1.5	87387	9	AC062037	AC062037 Homo sapi
C 426	19	1.5	9983	6	AX348459	AX348459 Sequence	C 499	19	1.5	90341	8	AC006232	AC006232 Arabidops
C 427	19	1.5	10114	1	U32812	U32812 Haemophilus	C 500	19	1.5	92127	2	AC112953	AC112953 Continuation (4 of
C 428	19	1.5	10842	1	AE004821	AE004821 Pseudomon	C 501	19	1.5	92203	2	AP001320	AP001320 Homo sapi
C 429	19	1.5	13216	6	AX344172	AX344172 Sequence	C 502	19	1.5	92280	5	AX625520	AX625520 Zebrafish
C 430	19	1.5	13249	6	AX345018	AX345018 Sequence	C 503	19	1.5	92896	9	AL603908	AL603908 Human DNA
C 431	19	1.5	13249	6	AX348563	AX348563 Sequence	C 504	19	1.5	93791	2	AC138073	AC138073 Homo sapi
C 432	19	1.5	13684	3	AE001403	AE001403 Plasmodiu	C 505	19	1.5	94128	2	AC131393	AC131393 Lotus cor
C 433	19	1.5	14939	1	AE013057	AE013057 Thermoana	C 506	19	1.5	95559	8	AP006538	AP006538 Homo sapi
C 434	19	1.5	17211	6	AX345556	AX345556 Sequence	C 507	19	1.5	95567	8	AC091232	AC091232 Homo sapi
C 435	19	1.5	17211	6	AX346631	AX346631 Sequence	C 508	19	1.5	96880	5	AC107164	AC107164 Danio rer
C 436	19	1.5	22049	2	AC091737	AC091737 Homo sapi	C 509	19	1.5	97191	2	CR752661	CR752661 Homo sapi
C 437	19	1.5	24993	1	AY343331	AY343331 Ehrlichia	C 510	19	1.5	97331	2	AX927213	AX927213 Danio rer
C 438	19	1.5	25595	10	BX322599	BX322599 Mouse DNA	C 511	19	1.5	98645	3	AC005267	AC005267 Drosophill
C 439	19	1.5	27195	3	U55854	U55854 Caenorhabdi	C 512	19	1.5	98645	3	AC005267	AC005267 Drosophill
C 440	19	1.5	28420	9	AL162391	AL162391 Human DNA	C 513	19	1.5	99799	9	AL445184	AL445184 Human DNA
C 441	19	1.5	29605	8	AB020747	AB020747 Arabidops	C 514	19	1.5	99973	9	AL161797	AL161797 Homo sapi
C 442	19	1.5	30758	2	CEH06104	Z94122 Caenorhabdi	C 515	19	1.5	100571	9	AL161797	AL161797 Human DNA
C 443	19	1.5	30961	3	CEH36H2	Z81078 Caenorhabdi	C 516	19	1.5	101539	8	ATT5P19	ATT5P19 Arabidops
C 444	19	1.5	31427	10	BX322602	BX322602 Mouse DNA	C 517	19	1.5	101542	9	AL391216	AL391216 Human DNA
C 445	19	1.5	31831	3	AC005457	AC005457 Drosophill	C 518	19	1.5	102393	8	AC000107	AC000107 Genomic s
C 446	19	1.5	32028	3	AF000197	AF000197 Caenorhab	C 519	19	1.5	104881	9	CR759848	CR759848 Human DNA
C 447	19	1.5	32298	8	AP006641	AP006641 Lotus cor	C 520	19	1.5	104913	9	AC004242	AC004242 Homo sapi
C 448	19	1.5	32953	6	AX458490	AX458490 Sequence	C 521	19	1.5	106100	2	AC138099	AC138099 Mus muscu
C 449	19	1.5	33480	3	AB010261	AB010261 Drosophill	C 522	19	1.5	106226	2	AX927090	AX927090 Homo sapi
C 450	19	1.5	33602	9	U73643	U73643 Human Chrom	C 523	19	1.5	107447	10	BX546504	BX546504 Mouse DNA
C 451	19	1.5	34754	4	SSC494748	A494748 Sus scrof	C 524	19	1.5	107795	8	AC149482	AC149482 Populus b
C 452	19	1.5	35585	9	AF108694	AF108694 Homo sapi	C 525	19	1.5	107959	8	AX649576	AX649576 Zebrafish
C 453	19	1.5	38092	3	AF016429	AF016429 Caenorhab	C 526	19	1.5	109081	5	CNS07YPH	AL713936 Oryza sat
C 454	19	1.5	39184	3	AL606823	AL606823 Human DNA	C 527	19	1.5	109326	9	AX927168	AX927168 Human DNA
C 455	19	1.5	39373	5	AC006625	AC006625 Caenorhab	C 528	19	1.5	109757	9	AL162379	AL162379 Human DNA
C 456	19	1.5	40832	5	AX539344	AX539344 Zebrafish	C 529	19	1.5	110000	1	AP006628	AP006628 Continuation (5 of
C 457	19	1.5					C 530	19	1.5	110000	1	AP006628	AP006628 Continuation (6 of

C 531	19	1.5 110000	1	AP006628_7	Continuation (8 of	C 604	19	1.5 139771	9	CR753846	CR753846 Human DNA
C 532	19	1.5 110000	2	AX908798_18	Continuation (19 of	C 605	19	1.5 139904	9	HS97D16	AL0091779 Human DNA
C 533	19	1.5 110000	1	AC096119_1	Continuation (2 of	C 606	19	1.5 139961	9	AL606477	AL606477 Human DNA
C 534	19	1.5 110000	2	AC096119_2	Continuation (3 of	C 607	19	1.5 140478	2	AC138988	AC138988 Homo sapi
C 535	19	1.5 110000	2	AC096343_0	AC096343 Rattus no	C 608	19	1.5 140643	8	AP003888	AP003888 Oryza sat
C 536	19	1.5 110000	2	AC118915_2	Continuation (3 of	C 609	19	1.5 140936	2	AC008630	AC008630 Homo sapi
C 537	19	1.5 110000	2	AY294423_0	AY294423 Mus muscu	C 610	19	1.5 141156	9	AL360093	AL360093 Human DNA
C 538	19	1.5 110000	2	BX842678_1	Continuation (2 of	C 611	19	1.5 141439	10	AC105403	AC105403 Mus muscu
C 539	19	1.5 110000	2	BX957276_1	Continuation (2 of	C 612	19	1.5 142159	10	AC092479	AC092479 Mus muscu
C 540	19	1.5 110000	2	BX957283_0	BX957283 Danio rer	C 613	19	1.5 142673	10	AC122045	AC122045 Mus muscu
C 541	19	1.5 110000	2	BX957283_4	Continuation (2 of	C 614	19	1.5 143280	10	AL845285	AL845285 Mouse DNA
C 542	19	1.5 110000	2	CEY10588_4	Continuation (5 of	C 615	19	1.5 144474	9	AF051934	AF051934 Homo sapi
C 543	19	1.5 110000	3	AC116984_1	Continuation (2 of	C 616	19	1.5 144848	2	AC140069	AC140069 Mus muscu
C 544	19	1.5 110000	3	AC116984_2	Continuation (3 of	C 617	19	1.5 145416	5	AC147826	AC147826 Xenopus t
C 545	19	1.5 110000	6	BD426631_13	Continuation (14 of	C 618	19	1.5 145496	8	AC144342	AC144342 Medicago
C 546	19	1.5 110000	6	BD426631_14	Continuation (15 of	C 619	19	1.5 145962	2	AC144373	AC144373 Felis cat
C 547	19	1.5 110000	6	AR274513_13	Continuation (14 of	C 620	19	1.5 146060	10	AC115849	AC115849 Mus muscu
C 548	19	1.5 110000	6	AR274513_14	Continuation (15 of	C 621	19	1.5 146259	10	AC102847	AC102847 Mus muscu
C 549	19	1.5 110000	6	AR541453_13	Continuation (14 of	C 622	19	1.5 146454	14	AF410153	AF410153 Swinepox
C 550	19	1.5 110000	6	AR541453_14	Continuation (15 of	C 623	19	1.5 146502	9	AC116456	AC116456 Homo sapi
C 551	19	1.5 110243	8	AC146945	AC146945 Oryza sat	C 624	19	1.5 147086	3	AC009182	AC009182 Drosophila
C 552	19	1.5 111130	8	AC144516	AC144516 Medicago	C 625	19	1.5 147118	9	AC007513	AC007513 Homo sapi
C 553	19	1.5 111271	9	AC011393	AC011393 Homo sapi	C 626	19	1.5 147318	9	AC138996	AC138996 Homo sapi
C 554	19	1.5 112027	9	AC007006	AC007006 Homo sapi	C 627	19	1.5 147369	2	AC147689	AC147689 Cercopith
C 555	19	1.5 113851	2	AP000571	AP000571 Homo sapi	C 628	19	1.5 147440	9	AP003971	AP003971 Homo sapi
C 556	19	1.5 114495	9	AC011743	AC011743 Homo sapi	C 629	19	1.5 147557	9	AL662789	AL662789 Human DNA
C 557	19	1.5 115177	9	AL583836	AL583836 Human DNA	C 630	19	1.5 148235	9	AC146372	AC146372 Pan trogl
C 558	19	1.5 115306	9	AC008615	AC008615 Homo sapi	C 631	19	1.5 148860	8	OSJN00093	AL607123 Mouse DNA
C 559	19	1.5 115808	8	AC104713	Oryza sat	C 632	19	1.5 149369	10	AL607123	AL607123 Mouse DNA
C 560	19	1.5 116180	2	AC133296	Rattus no	C 633	19	1.5 149599	14	AY077835	AY077835 Goatpox v
C 561	19	1.5 116960	9	AC135632	Continuation (14 of	C 634	19	1.5 149662	14	AY077834	AY077834 Sheepox
C 562	19	1.5 116989	8	AC137819	Medicago	C 635	19	1.5 149723	14	AY077836	AY077836 Goatpox v
C 563	19	1.5 117303	9	AL773543	Continuation (14 of	C 636	19	1.5 149788	8	ATT6G321	AL589883 Arabidops
C 564	19	1.5 117645	8	AC144644	Medicago	C 637	19	1.5 149955	14	AY077832	AY077832 Sheepox
C 565	19	1.5 118610	2	AC002344	Homo sapi	C 638	19	1.5 150029	10	BX890639	BX890639 Mouse DNA
C 566	19	1.5 119077	2	AC147494	Carollia	C 639	19	1.5 150057	14	AY077833	AY077833 Sheepox
C 567	19	1.5 119660	2	AC149578	Continuation (14 of	C 640	19	1.5 150434	9	AL138784	AL138784 Human DNA
C 568	19	1.5 119706	2	AC148352	Continuation (14 of	C 641	19	1.5 150509	14	AF409138	AF409138 Lumby ski
C 569	19	1.5 119765	2	AC090705	Homo sapi	C 642	19	1.5 150773	14	AF325528	AF325528 Lumby ski
C 570	19	1.5 120037	2	AC147766	Dasypus n	C 643	19	1.5 150793	14	AF409137	AF409137 Lumby ski
C 571	19	1.5 121059	9	AL138918	Human DNA	C 644	19	1.5 151254	5	BX649371	BX649371 Zebrafish
C 572	19	1.5 121806	2	CR759819	Danio rer	C 645	19	1.5 151329	9	BS000541	BS000541 Pan trogl
C 573	19	1.5 122489	2	AC027339	Continuation (14 of	C 646	19	1.5 151349	2	AL353647	AL353647 Homo sapi
C 574	19	1.5 122880	9	AC016711	Homo sapi	C 647	19	1.5 151517	9	AC105230	AC105230 Homo sapi
C 575	19	1.5 122948	9	AL355798	Human DNA	C 648	19	1.5 151519	9	HS23114	AL022719 Human DNA
C 576	19	1.5 123323	9	AP001816	Homo sapi	C 649	19	1.5 151551	2	AC021291	AC021291 Homo sapi
C 577	19	1.5 124372	10	AL731808	Mouse DNA	C 650	19	1.5 151744	10	AL772406	AL772406 Mouse DNA
C 578	19	1.5 124787	9	BX248406	Human DNA	C 651	19	1.5 152106	10	AC119857	AC119857 Mus muscu
C 579	19	1.5 125381	4	AY644517	AY644517 Bos tauru	C 652	19	1.5 152264	8	OSJN00090	AL606685 Oryza sat
C 580	19	1.5 125674	2	AP002018	Homo sapi	C 653	19	1.5 152484	2	AC067992	AC067992 Homo sapi
C 581	19	1.5 126141	2	AC131240	Medicago	C 654	19	1.5 152517	9	AC009468	AC009468 Homo sapi
C 582	19	1.5 127298	10	BX005130	Mouse DNA	C 655	19	1.5 152649	2	CR356241	CR356241 Danio rer
C 583	19	1.5 128024	2	AC147955	Dasypus n	C 656	19	1.5 152695	2	AP005152	AP005152 Oryza sat
C 584	19	1.5 128152	9	AC113347	Continuation (14 of	C 657	19	1.5 152877	9	AC107399	AC107399 Homo sapi
C 585	19	1.5 129467	9	HSJ601K24	Continuation (14 of	C 658	19	1.5 153031	2	AC105458	AC105458 Felis cat
C 586	19	1.5 131259	2	AP001805	Homo sapi	C 659	19	1.5 153078	9	AP002782	AP002782 Homo sapi
C 587	19	1.5 131727	9	AC093850	Homo sapi	C 660	19	1.5 153616	2	AC149696	AC149696 Bos tauru
C 588	19	1.5 131885	1	CR628338	Legionell	C 661	19	1.5 154279	2	AC151077	AC151077 Bos tauru
C 589	19	1.5 132699	8	AC006917	Genomic a	C 662	19	1.5 154373	2	AC146803	AC146803 Xenopus t
C 590	19	1.5 132742	2	AC145412	Bos tauru	C 663	19	1.5 154713	9	AC013361	AC013361 Homo sapi
C 591	19	1.5 133330	10	AL928871	Mouse DNA	C 664	19	1.5 154907	9	AC079631	AC079631 Homo sapi
C 592	19	1.5 134054	9	AC026730	Homo sapi	C 665	19	1.5 155100	2	AC102514	AC102514 Mus muscu
C 593	19	1.5 135423	9	AC112195	Homo sapi	C 666	19	1.5 155437	2	AC116872	AC116872 Mus muscu
C 594	19	1.5 135679	2	AC138993	Homo sapi	C 667	19	1.5 155535	8	AP016909	AP016909 Homo sapi
C 595	19	1.5 135998	9	AC104021	Homo sapi	C 668	19	1.5 155542	8	AP004662	AP004662 Oryza sat
C 596	19	1.5 136511	9	AC117424	Homo sapi	C 669	19	1.5 156022	2	AC110560	AC110560 Mus muscu
C 597	19	1.5 136799	8	AP004564	Oryza sat	C 670	19	1.5 156568	9	AC097474	AC097474 Homo sapi
C 598	19	1.5 137113	9	AC076966	Homo sapi	C 671	19	1.5 157042	2	CR759826	CR759826 Danio rer
C 599	19	1.5 137325	8	AP005397	Oryza sat	C 672	19	1.5 157330	9	AC025207	AC025207 Homo sapi
C 600	19	1.5 137525	8	AP005397	Oryza sat	C 673	19	1.5 157530	9	AC146247	AC146247 Pan trogl
C 601	19	1.5 138391	2	AL357733	Homo sapi	C 674	19	1.5 157556	2	CR354606	CR354606 Danio rer
C 602	19	1.5 139015	2	AC087581	Homo sapi	C 675	19	1.5 157631	8	AP005545	AP005545 Oryza sat
C 603	19	1.5 139032	9	AC010656	Homo sapi	C 676	19	1.5 157875	9	HS272156	AL023754 Human DNA

677	19	1.5	158180	9	AC100781	AC100781 Homo sapi	c 750	19	1.5	172750	2	AC119495	AC119495 Rattus no
678	19	1.5	158500	10	AC131983	Mus muscu	751	19	1.5	173639	10	AL606921	AL606921 Mouse DNA
c 679	19	1.5	158515	2	CR318625	Danio rer	752	19	1.5	174313	9	AC024060	AC024060 Homo sapi
c 680	19	1.5	158563	2	AC150484	Bos tauru	c 753	19	1.5	174313	9	AC024060	AC024060 Homo sapi
c 681	19	1.5	158830	10	AC131740	Mus muscu	754	19	1.5	174896	2	AC022012	AC022012 Homo sapi
c 682	19	1.5	159317	2	AC147520	Otolemur	c 755	19	1.5	174927	9	AC010881	AC010881 Homo sapi
c 683	19	1.5	159333	5	BX004980	Zebraphish	c 756	19	1.5	175003	10	AC115900	AC115900 Mus muscu
c 684	19	1.5	159658	9	AC104062	Homo sapi	c 757	19	1.5	175152	2	AC0101965	AC0101965 Mus muscu
c 685	19	1.5	159705	2	AC021007	Homo sapi	c 758	19	1.5	175179	2	AC072025	AC072025 Homo sapi
c 686	19	1.5	159769	2	AC007337	Homo sapi	c 759	19	1.5	175245	10	AC129188	AC129188 Mus muscu
c 687	19	1.5	159792	10	AC139182	Mus muscu	c 760	19	1.5	175467	2	AC079215	AC079215 Homo sapi
c 688	19	1.5	160012	5	BX005254	Zebraphish	761	19	1.5	175809	10	AC099629	AC099629 Mus muscu
c 689	19	1.5	160157	2	AC091294	Mus muscu	762	19	1.5	175916	9	AC018842	AC018842 Homo sapi
c 690	19	1.5	160462	2	CR318609	Danio rer	c 763	19	1.5	175977	9	AC027691	AC027691 Homo sapi
c 691	19	1.5	160511	2	AC022786	Homo sapi	c 764	19	1.5	176051	9	AC022008	AC022008 Homo sapi
c 692	19	1.5	160866	2	AC120170	Mus muscu	765	19	1.5	176104	9	AC007163	AC007163 Homo sapi
c 693	19	1.5	160713	2	AC007174	Homo sapi	c 766	19	1.5	176594	2	AC119325	AC119325 Rattus no
c 694	19	1.5	161099	10	AC132624	Mus muscu	c 767	19	1.5	176983	2	AC120820	AC120820 Rattus no
c 695	19	1.5	161300	2	AC026169	Homo sapi	c 768	19	1.5	177087	2	AC093687	AC093687 Homo sapi
c 696	19	1.5	161377	2	AC141827	Apis mell	c 769	19	1.5	177097	2	AP001569	AP001569 Homo sapi
c 697	19	1.5	161739	9	AC013302	Homo sapi	c 770	19	1.5	177245	9	AC073856	AC073856 Homo sapi
c 698	19	1.5	162030	2	AC018601	Homo sapi	c 771	19	1.5	177363	9	AC073856	AC073856 Homo sapi
c 699	19	1.5	162270	2	AC140470	Homo sapi	c 772	19	1.5	177623	2	BX927353	DX927353 Danio rer
c 700	19	1.5	162598	2	AC079091	Homo sapi	c 773	19	1.5	177849	2	AC125624	AC125624 Homo sapi
701	19	1.5	162709	10	AL807237	Mouse DNA	c 774	19	1.5	177929	9	AC125624	AC125624 Homo sapi
c 702	19	1.5	163025	2	AC024063	Homo sapi	775	19	1.5	178046	2	AC137949	AC137949 Mus muscu
c 703	19	1.5	163126	9	AC110012	Homo sapi	c 776	19	1.5	178207	9	AC140172	AC140172 Homo sapi
c 704	19	1.5	163228	9	AC009039	Homo sapi	c 777	19	1.5	178317	9	AC012081	AC012081 Homo sapi
c 705	19	1.5	163331	9	AL162575	Human DNA	c 778	19	1.5	178420	9	AP003548	AP003548 Homo sapi
c 706	19	1.5	163385	10	AL162575	Human DNA	c 779	19	1.5	178534	2	AC024956	AC024956 Homo sapi
707	19	1.5	163814	5	AL953863	Zebraphish	c 780	19	1.5	178800	2	AC149166	AC149166 Otolemur
708	19	1.5	163926	9	AC096634	Homo sapi	781	19	1.5	178824	5	BX323024	BX323024 Zebraphish
709	19	1.5	164190	9	AC025448	Homo sapi	c 782	19	1.5	179042	10	AC119905	AC119905 Mus muscu
710	19	1.5	164575	2	AC024736	Homo sapi	c 783	19	1.5	179547	2	CR762491	CR762491 Danio rer
711	19	1.5	164781	9	AC104419	Homo sapi	c 784	19	1.5	179583	2	CR788312	CR788312 Danio rer
712	19	1.5	164808	2	AC023969	Homo sapi	c 785	19	1.5	179590	10	AC113293	AC113293 Mus muscu
713	19	1.5	165349	10	AC134870	Mus muscu	c 786	19	1.5	179681	2	AC013785	AC013785 Homo sapi
714	19	1.5	165789	2	AC119054	Rattus no	c 787	19	1.5	179726	9	AC007052	AC007052 Homo sapi
715	19	1.5	165992	2	AC149707	Bos tauru	c 788	19	1.5	179814	9	AC068314	AC068314 Homo sapi
716	19	1.5	166351	10	AC118747	Mus muscu	789	19	1.5	179934	2	AC128793	AC128793 Rattus no
717	19	1.5	166398	9	AC090568	Homo sapi	c 790	19	1.5	180000	2	AC006428	AC006428 Homo sapi
c 718	19	1.5	166522	2	AL357072	Homo sapi	c 791	19	1.5	180311	9	AC140086	AC140086 Cercopit
c 719	19	1.5	166644	9	AC0011352	Homo sapi	c 792	19	1.5	180401	9	AC007489	AC007489 Homo sapi
c 720	19	1.5	166743	9	AC005598	Homo sapi	c 793	19	1.5	180425	2	CR381605	CR381605 Danio rer
c 721	19	1.5	166847	2	BXS06C7S	Human chr	c 794	19	1.5	180861	2	CR391970	CR391970 Homo sapi
c 722	19	1.5	167207	2	BN571889	Mus muscu	c 795	19	1.5	180973	9	AC097462	AC097462 Homo sapi
c 723	19	1.5	167292	2	AC068365	Homo sapi	c 796	19	1.5	181258	2	AC147355	AC147355 Xenopus t
c 724	19	1.5	167386	2	AC013704	Homo sapi	c 797	19	1.5	181963	2	AC102659	AC102659 Mus muscu
c 725	19	1.5	167596	2	AC013177	Drosophila	c 798	19	1.5	182028	2	AC102442	AC102442 Mus muscu
c 726	19	1.5	169048	9	AL158011	Human DNA	c 799	19	1.5	182411	2	AC090408	AC090408 Homo sapi
c 727	19	1.5	169063	9	AC098677	Homo sapi	c 800	19	1.5	182520	10	AL627392	AL627392 Mouse DNA
c 728	19	1.5	169199	9	AC018812	Homo sapi	c 801	19	1.5	182555	9	AC068531	AC068531 Homo sapi
c 729	19	1.5	169465	10	AL606976	Mouse DNA	c 802	19	1.5	182555	9	AC007282	AC007282 Homo sapi
c 730	19	1.5	169686	2	BX465864	Danio rer	c 803	19	1.5	182696	10	AC126424	AC126424 Mus muscu
c 731	19	1.5	169891	2	AC114479	Homo sapi	c 804	19	1.5	182743	2	AC046132	AC046132 Homo sapi
c 732	19	1.5	170022	2	AC021559	Homo sapi	c 805	19	1.5	183007	3	AC007357	AC007357 Drosophila
c 733	19	1.5	170254	10	AC137678	Mus muscu	c 806	19	1.5	183330	9	AC011372	AC011372 Homo sapi
734	19	1.5	170452	8	AF152364	Homo sapi	c 807	19	1.5	183344	9	AC147112	AC147112 Pan trogl
735	19	1.5	170523	8	AP004764	Oryza sat	c 808	19	1.5	183477	2	BX957239	BX957239 Danio rer
c 736	19	1.5	170624	9	AC006031	Homo sapi	c 809	19	1.5	183638	10	AC132345	AC132345 Mus muscu
737	19	1.5	170624	2	AC144620	Mus muscu	c 810	19	1.5	184010	2	AC150025	AC150025 Papio anu
738	19	1.5	170631	9	AC020751	Homo sapi	c 811	19	1.5	184562	2	AC102371	AC102371 Mus muscu
739	19	1.5	170750	2	AC103898	Bos tauru	c 812	19	1.5	184591	9	AL365443	AL365443 Human DNA
740	19	1.5	170759	2	AC138994	Homo sapi	c 813	19	1.5	184989	2	AC113392	AC113392 Homo sapi
741	19	1.5	170755	9	AL589745	Human DNA	c 814	19	1.5	185138	10	AL845172	AL845172 Mouse DNA
742	19	1.5	170807	9	AC023812	Homo sapi	c 815	19	1.5	185209	9	AC145856	AC145856 Pan trogl
743	19	1.5	171019	3	AC099031	Drosophila	c 816	19	1.5	185242	10	AC115734	AC115734 Mus muscu
744	19	1.5	171185	2	AP002424	Homo sapi	c 817	19	1.5	185242	10	AC124597	AC124597 Mus muscu
745	19	1.5	171747	2	AC087678	Homo sapi	c 818	19	1.5	185806	9	AC147049	AC147049 Pan trogl
746	19	1.5	172155	2	AC018653	Homo sapi	c 819	19	1.5	186058	2	AC112245	AC112245 Homo sapi
747	19	1.5	172285	2	CR381538	Danio rer	c 820	19	1.5	186064	2	AC020752	AC020752 Homo sapi
c 748	19	1.5	172346	2	AC016349	Homo sapi	c 821	19	1.5	186561	2	AC102764	AC102764 Mus muscu
c 749	19	1.5	172748	9	AC058871	Homo sapi	c 822	19	1.5	186608	9	AC093274	AC093274 Homo sapi

C 823	1.5	186670	9	AC080188	AC080188 Homo sapi	C 896	1.5	200290	2	EX000350	EX000350 Mus muscu
C 824	1.5	186742	5	AL928736	AL928736 Zebrafish	C 897	1.5	200774	2	AP001592	AP001592 Homo sapi
C 825	1.5	186773	10	AC135356	AC135356 Mus muscu	C 898	1.5	201162	10	AC069561	AC069561 Mus muscu
C 826	1.5	186839	10	AC147152	AC147152 Mus muscu	C 899	1.5	201359	2	AC128640	AC128640 Homo sapi
C 827	1.5	187059	2	CR318671	CR318671 Danio rer	C 900	1.5	201611	2	AC021185	AC021185 Homo sapi
C 828	1.5	187127	2	AC119942	AC119942 Mus muscu	C 901	1.5	201557	2	AC006840	AC006840 Homo sapi
C 829	1.5	187241	2	AL772146	AL772146 Zebrafish	C 902	1.5	202396	2	AC140011	AC140011 Homo sapi
C 830	1.5	187417	2	AC105154	AC105154 Mus muscu	C 903	1.5	202412	10	AL670673	AL670673 Mouse DNA
C 831	1.5	187595	2	AC138878	AC138878 Homo sapi	C 904	1.5	202469	2	AC099607	AC099607 Mus muscu
C 832	1.5	187727	2	AC064806	AC064806 Homo sapi	C 905	1.5	202676	2	AC129140	AC129140 Rattus no
C 833	1.5	187898	2	AC016715	AC016715 Homo sapi	C 906	1.5	203061	9	AC105054	AC105054 Homo sapi
C 834	1.5	188084	2	AC124153	AC124153 Danio rer	C 907	1.5	203120	2	AC148499	AC148499 Callithri
C 835	1.5	188275	2	AC150889	AC150889 Bos tauru	C 908	1.5	203204	9	AC092052	AC092052 Homo sapi
C 836	1.5	188476	10	AL773549	AL773549 Mouse DNA	C 909	1.5	203399	2	AC121559	AC121559 Mus muscu
C 837	1.5	188484	10	AC122749	AC122749 Mus muscu	C 910	1.5	204632	2	AC097583	AC097583 Rattus no
C 838	1.5	188700	10	AC121525	AC121525 Mus muscu	C 911	1.5	204988	2	CR293501	CR293501 Danio rer
C 839	1.5	188958	10	AL844605	AL844605 Mouse DNA	C 912	1.5	205033	2	AC111383	AC111383 Rattus no
C 840	1.5	188963	2	CR762411	CR762411 Danio rer	C 913	1.5	205152	9	AP002985	AP002985 Homo sapi
C 841	1.5	189412	2	AC023532	AC023532 Homo sapi	C 914	1.5	206332	2	AC116270	AC116270 Rattus no
C 842	1.5	189466	9	AC018887	AC018887 Homo sapi	C 915	1.5	206474	2	AC129671	AC129671 Rattus no
C 843	1.5	189563	5	EX000364	EX000364 Zebrafish	C 916	1.5	207365	2	AC093653	AC093653 Homo sapi
C 844	1.5	189718	2	AC128388	AC128388 Rattus no	C 917	1.5	207802	10	AC127326	AC127326 Mus muscu
C 845	1.5	190015	2	CR376768	CR376768 Danio rer	C 918	1.5	208059	5	AL935193	AL935193 Zebrafish
C 846	1.5	190015	2	AC140056	AC140056 Rattus no	C 919	1.5	208499	2	CR376859	CR376859 Danio rer
C 847	1.5	190220	9	AC142323	AC142323 Pan trogl	C 920	1.5	208726	9	AP006287	AP006287 Homo sapi
C 848	1.5	190379	2	AC022386	AC022386 Homo sapi	C 921	1.5	209176	2	AC110516	AC110516 Mus muscu
C 849	1.5	190577	2	AC073721	AC073721 Mus muscu	C 922	1.5	209848	2	CR388132	CR388132 Danio rer
C 850	1.5	190702	5	EX119962	EX119962 Zebrafish	C 923	1.5	210161	2	EX470245	EX470245 Danio rer
C 851	1.5	190816	2	AC116446	AC116446 Mus muscu	C 924	1.5	210640	5	EX293994	EX293994 Zebrafish
C 852	1.5	191355	2	AC151028	AC151028 Callithri	C 925	1.5	210772	2	AC121080	AC121080 Mus muscu
C 853	1.5	191702	2	AC148257	AC148257 Callithri	C 926	1.5	211076	2	AC012214	AC012214 Homo sapi
C 854	1.5	191896	9	AC009481	AC009481 Homo sapi	C 927	1.5	211367	2	AC114176	AC114176 Rattus no
C 855	1.5	191900	2	AC121041	AC121041 Rattus no	C 928	1.5	211540	2	AC117161	AC117161 Rattus no
C 856	1.5	192020	10	AC140303	AC140303 Mus muscu	C 929	1.5	212472	10	AL731838	AL731838 Mouse DNA
C 857	1.5	192046	9	AC008083	AC008083 Homo sapi	C 930	1.5	212954	9	AC144376	AC144376 Pan trogl
C 858	1.5	192685	2	AC141749	AC141749 Apis mell	C 931	1.5	214337	2	EX557231	EX557231 Danio rer
C 859	1.5	193115	2	AC134667	AC134667 Mus muscu	C 932	1.5	214701	2	AP001447	AP001447 Homo sapi
C 860	1.5	193290	2	CR396590	CR396590 Danio rer	C 933	1.5	215286	9	AC130650	AC130650 Homo sapi
C 861	1.5	193353	2	AC113442	AC113442 Mus muscu	C 934	1.5	215286	9	HSU91322	U91322 Human chrom
C 862	1.5	193785	10	AL772201	AL772201 Mouse DNA	C 935	1.5	215385	9	AF216808	AF216808 Homo sapi
C 863	1.5	193849	2	AC084829	AC084829 Mus muscu	C 936	1.5	215521	2	AC126103	AC126103 Rattus no
C 864	1.5	193917	10	AC130817	AC130817 Mus muscu	C 937	1.5	215772	2	AC132188	AC132188 Pan trogl
C 865	1.5	193978	2	AC123751	AC123751 Mus muscu	C 938	1.5	216041	2	AC130019	AC130019 Rattus no
C 866	1.5	194063	2	AC118643	AC118643 Mus muscu	C 939	1.5	216334	9	AC011132	AC011132 Homo sapi
C 867	1.5	194183	2	AC026010	AC026010 Homo sapi	C 940	1.5	217145	2	AC135513	AC135513 Rattus no
C 868	1.5	194553	9	AC097536	AC097536 Homo sapi	C 941	1.5	217707	5	EX248129	EX248129 Zebrafish
C 869	1.5	194881	2	AC018967	AC018967 Homo sapi	C 942	1.5	220068	10	AC125332	AC125332 Mus muscu
C 870	1.5	195186	2	AC149309	AC149309 Zea mays	C 943	1.5	220379	2	AC136271	AC136271 Rattus no
C 871	1.5	195451	2	AC140138	AC140138 Homo sapi	C 944	1.5	221043	10	AC148991	AC148991 Mus muscu
C 872	1.5	195932	2	AC021710	AC021710 Homo sapi	C 945	1.5	221963	2	AC024363	AC024363 Mus muscu
C 873	1.5	195936	2	CR847988	CR847988 Danio rer	C 946	1.5	218552	10	AC118607	AC118607 Mus muscu
C 874	1.5	196661	2	AC133197	AC133197 Mus muscu	C 947	1.5	219461	2	EX842685	EX842685 Danio rer
C 875	1.5	196759	2	AC113902	AC113902 Rattus no	C 948	1.5	220068	10	AC125332	AC125332 Mus muscu
C 876	1.5	196803	10	AC132461	AC132461 Mus muscu	C 949	1.5	220379	2	AC136271	AC136271 Rattus no
C 877	1.5	196869	2	AC087535	AC087535 Homo sapi	C 950	1.5	221043	10	AC148991	AC148991 Mus muscu
C 878	1.5	196891	2	AC136316	AC136316 Mus muscu	C 951	1.5	221091	9	AP003390	AP003390 Homo sapi
C 879	1.5	196898	2	AC123721	AC123721 Mus muscu	C 952	1.5	221618	10	AC092530	AC092530 Rattus no
C 880	1.5	196905	2	AC121641	AC121641 Rattus no	C 953	1.5	222041	2	AC116093	AC116093 Rattus no
C 881	1.5	196989	9	BS000196	BS000196 Pan trogl	C 954	1.5	222161	2	AC147935	AC147935 Ocolemur
C 882	1.5	197162	2	AC141596	AC141596 Homo sapi	C 955	1.5	222520	2	AC120336	AC120336 Rattus no
C 883	1.5	197384	2	AC129568	AC129568 Mus muscu	C 956	1.5	222801	2	AC093951	AC093951 Rattus no
C 884	1.5	197785	10	AL844865	AL844865 Mouse DNA	C 957	1.5	222825	2	CR628387	CR628387 Danio rer
C 885	1.5	198056	2	AC078951	AC078951 Homo sapi	C 958	1.5	222885	2	AL158816	AL158816 Homo sapi
C 886	1.5	198472	2	AC144586	AC144586 Homo sapi	C 959	1.5	223207	2	AC112065	AC112065 Rattus no
C 887	1.5	198491	2	AC022249	AC022249 Homo sapi	C 960	1.5	223606	2	AC013479	AC013479 Homo sapi
C 888	1.5	198838	2	AC093647	AC093647 Homo sapi	C 961	1.5	223816	5	EX537162	EX537162 Zebrafish
C 889	1.5	198946	10	AL844548	AL844548 Mouse DNA	C 962	1.5	224100	2	AC119827	AC119827 Mus muscu
C 890	1.5	199245	2	AC138622	AC138622 Mus muscu	C 963	1.5	224749	2	AC107594	AC107594 Rattus no
C 891	1.5	199569	9	AC016250	AC016250 Homo sapi	C 964	1.5	226664	2	AC139775	AC139775 Homo sapi
C 892	1.5	199885	10	AC147376	AC147376 Mus muscu	C 965	1.5	226936	2	AC132654	AC132654 Rattus no
C 893	1.5	200019	2	AC024111	AC024111 Mus muscu	C 966	1.5	227294	2	CR847824	CR847824 Danio rer
C 894	1.5	200045	2	AC122678	AC122678 Rattus no	C 967	1.5	227781	2	AC129745	AC129745 Rattus no
C 895	1.5	200231	2	AC123230	AC123230 Rattus no	C 968	1.5	227783	2	AC127601	AC127601 Rattus no



c 969 1.5 228044 5 BX682552  
970 1.5 228298 2 AY657026  
971 1.5 228428 2 AC099217  
c 972 1.5 228458 2 AC097670  
973 1.5 228732 2 CR391929  
974 1.5 229076 2 AC107142  
c 975 1.5 229278 10 AL672063  
976 1.5 229289 10 AL603842  
977 1.5 229480 5 AL929150  
978 1.5 229533 2 AC106382  
c 979 1.5 229731 2 AC103289  
980 1.5 229771 2 AC133813  
981 1.5 230391 2 AC136337  
c 982 1.5 230756 2 AC109986  
983 1.5 231675 10 AL591410  
984 1.5 232304 5 BX322616  
985 1.5 232385 2 AC099227  
986 1.5 232795 2 AC123149  
c 987 1.5 232874 2 AC114223  
988 1.5 233081 2 AC112460  
989 1.5 233134 2 AC120925  
990 1.5 233296 2 CR318636  
991 1.5 234428 2 AC113793  
c 992 1.5 234848 2 AC109367  
993 1.5 234983 2 AC099168  
c 994 1.5 235267 2 CR392024  
995 1.5 235405 10 AC101737  
996 1.5 235471 2 AC135649  
997 1.5 235720 2 AC120712  
c 998 1.5 235757 2 AC121367  
999 1.5 236716 2 AC098920  
c1000 1.5 236717 2 AC107440

## ALIGNMENTS

RESULT 1  
AR411820 1274 bp DNA linear PAT 18-DBC-2003  
LOCUS  
DEFINITION Sequence 1 from patent US 6638734.  
ACCESSION AR411820  
VERSION AR411820.1 GI:40164258  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 1274)  
Reed, J.C. and Matsuzawa, S.-i.  
Nucleic acid encoding proteins involved in protein degradation,  
products and methods related thereto  
Patent US 6638734-A 1 28-OCT-2003;  
JOURNAL Location/Qualifiers  
FEATURES  
source  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 100.0%; Score 1274; DB 6; Length 1274;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTAGTTGTTATGGTCCATTTTCTATTTAGCATTTATTTCTATGTAGTCTAT 60  
Db 1 TTTCTTAGTTGTTATGGTCCATTTTCTATTTAGCATTTATTTCTATGTAGTCTAT 60  
QY 61 CCAAGACGATTAAAGGAGTTCACATGTTCCTCGGAAACATTTTGAAGAGAGCTTATC 120  
Db 61 CCAAGACGATTAAAGGAGTTCACATGTTCCTCGGAAACATTTTGAAGAGAGCTTATC 120  
QY 121 CAGTGACAGATCCTTAATAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT 180  
Db 121 CAGTGACAGATCCTTAATAAGTGACATTCAGTGTAAATTTTATTTTAAATATCTTTT 180

QY 1261 CGGTAGTGGAGC 1274  
 Db 1261 CGGTAGTGGAGC 1274

RESULT 2  
 AX058082  
 LOCUS 1274 bp DNA linear PAT 17-JAN-2001  
 DEFINITION Sequence 1 from Patent WO0077207.  
 AX058082  
 ACCESSION  
 VERSION  
 AX058082.1 GI:12310662  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS  
 TITLE Nucleic acid encoding proteins involved in protein degradation,  
 products and methods related thereto  
 JOURNAL Patent: WO 0077207-A 1 21-DEC-2000;  
 The Burnham Institute (US)  
 FEATURES  
 Location/Qualifiers  
 1..1274  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 274..1170  
 /notes="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC22474.1"  
 /db\_xref="GI:12310663"  
 /translation="MVIILFLPPYVFISEMSQTATLPTGTSKPPSRVPLTGT  
 TASNDLALFECPFDVLPPILOCSGHLVCSNCRPKLTCPTCRPLGSLRLA  
 MEKVANSVLPCKYASSGCEITLPHTEKADHEELCEFRVPCPCGASCKWQGLDGH  
 MHLAHOXSIITLLOGEDIVELATDNLPCAVDWMQSCFGFHEMLVLEKQKDYDH  
 QQFALVQLIGTRKQENFAYRLNGLNHRRLTWEATPRSIHEGIAITMNSDCLVFD  
 TSAQLFAENGLNGLINVTISM"

ORIGIN  
 Query Match 100.0%; Score 1274; DB 6; Length 1274;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCTTTAGTTGTTATGTTCCATTTTCTATTTTAGCATTTATTTCTATGATGATCTAT 60  
 Db 1 TTTCTTTAGTTGTTATGTTCCATTTTCTATTTTAGCATTTATTTCTATGATGATCTAT 60

QY 61 CCAAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 120  
 Db 61 CCAAAGACGATTAAAGGAGTTCACATGTTTCCGGAACATTTTGAAGAGAGAGCTTATC 120

QY 121 CAGGTACAGATCTTAATAAGTGACATTCAGTGTAATTTTATTTTATTTTATCTTTT 180  
 Db 121 CAGGTACAGATCTTAATAAGTGACATTCAGTGTAATTTTATTTTATTTTATCTTTT 180

QY 181 TTAATCCTATTTTCTCTCTTTTCTCTCAGTAAATTTTGTATGAACTTTAAAGGACT 240  
 Db 181 TTAATCCTATTTTCTCTCTTTTCTCTCAGTAAATTTTGTATGAACTTTAAAGGACT 240

QY 241 TATGGCATGTAACATTTATTAAGTAAGTCAATCGTTATTAATTTTCTCTGCTT 300  
 Db 241 TATGGCATGTAACATTTATTAAGTAAGTCAATCGTTATTAATTTTCTCTGCTT 300

QY 301 CCTTATGTTATTTTTCAGAAATGAGCGCTCAGACTGCTACAGACTTACCTACCGTACC 360  
 Db 301 CCTTATGTTATTTTTCAGAAATGAGCGCTCAGACTGCTACAGACTTACCTACCGTACC 360

QY 361 TCGAAGTGTCCACATCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 Db 361 TCGAAGTGTCCACATCCAGAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTGACTATGTTGTTACCGCCCAATCTTT 480

Db 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTCTGCTTTGACTATGTTTACCGCCCAATCTTT 480  
 QY 481 CAATGTACAGATGGCCATCTTTGTTTGTAGCAACTGTGCGCCAAAGCTCAATGTTGTCCA 540  
 Db 481 CAATGTACAGATGGCCATCTTTGTTTGTAGCAACTGTGCGCCAAAGCTCAATGTTGTCCA 540

QY 541 ACTTGGCGGGGCCCTTTGGGATCCATTTGCAACTTGGCTATGAGAAAGTGGCTAATCA 600  
 Db 541 ACTTGGCGGGGCCCTTTGGGATCCATTTGCAACTTGGCTATGAGAAAGTGGCTAATCA 600

QY 601 GTACTTTTCCCTGTAAATATGCTCTCTGGATGTGAATAACTCTGCCACACACAGAA 660  
 Db 601 GTACTTTTCCCTGTAAATATGCTCTCTGGATGTGAATAACTCTGCCACACACAGAA 660

QY 661 AAAGCAGACCATGACAGCTCTGTGATGTTTAGGCCCTTATTTCTGTCGCTGCTGCTGCT 720  
 Db 661 AAAGCAGACCATGACAGCTCTGTGATGTTTAGGCCCTTATTTCTGTCGCTGCTGCTGCT 720

QY 721 TCTGTAAATGGAAGGCTCTCTGATGCTGTAATGCCCATCTGATGATACAGATAAG 780  
 Db 721 TCTGTAAATGGAAGGCTCTCTGATGCTGTAATGCCCATCTGATGATACAGATAAG 780

QY 781 TCCATTACACCCCTCAGGAGGAGGATATAGTTTTTCTTGTCTACACACATTAATCTTCT 840  
 Db 781 TCCATTACACCCCTCAGGAGGAGGATATAGTTTTTCTTGTCTACACACATTAATCTTCTCT 840

QY 841 GGTGCTGTTGACTGGGTGATGATGACAGTCTGTTTTTGGGCTTTTCACTTCACTGTTAGTCTTA 900  
 Db 841 GGTGCTGTTGACTGGGTGATGATGACAGTCTGTTTTTGGGCTTTTCACTTCACTGTTAGTCTTA 900

QY 901 GAGAAACAGGAAATAACGATGCTACAGAGCTTTTTCGCAATCTGTACAGCTGATAGGA 960  
 Db 901 GAGAAACAGGAAATAACGATGCTACAGAGCTTTTTCGCAATCTGTACAGCTGATAGGA 960

QY 961 ACACGCAAGCAGCTGAAATTTTGTCTACGACTTTGAGTAAATGCTATAGGCGACGA 1020  
 Db 961 ACACGCAAGCAGCTGAAATTTTGTCTACGACTTTGAGTAAATGCTATAGGCGACGA 1020

QY 1021 TTGACTTGGGAAAGCAGCTCTCTGATCTATTCATGAAGGAATTCGCAACAGCCATTTATGAAT 1080  
 Db 1021 TTGACTTGGGAAAGCAGCTCTCTGATCTATTCATGAAGGAATTCGCAACAGCCATTTATGAAT 1080

QY 1081 AGCAGCTGTCTAGTCTTTGACACAGCAGCTTTTTCGAGAAAATGGCAATTTA 1140  
 Db 1081 AGCAGCTGTCTAGTCTTTGACACAGCAGCTTTTTCGAGAAAATGGCAATTTA 1140

QY 1141 GGCATCAATGTAATCTATTTTCCATGTTGAAATGGCAATCAAAATTTTCTGCGCAGTGT 1200  
 Db 1141 GGCATCAATGTAATCTATTTTCCATGTTGAAATGGCAATCAAAATTTTCTGCGCAGTGT 1200

QY 1201 TTAATAACTTCAGTTTTCACAGAAAATAAGGACCCCATCTGCTGCGCAACCTTAAACTCTTT 1260  
 Db 1201 TTAATAACTTCAGTTTTCACAGAAAATAAGGACCCCATCTGCTGCGCAACCTTAAACTCTTT 1260

QY 1261 CGGTAGTGGAGC 1274  
 Db 1261 CGGTAGTGGAGC 1274

RESULT 3  
 CQ491092  
 LOCUS 2924 bp DNA linear PAT 30-JAN-2004  
 DEFINITION Sequence 22959 from Patent WO0160860.  
 CQ491092  
 ACCESSION  
 VERSION CQ491092.1 GI:41456711  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Schlegel, R., Endege, W.O. and Mohahan, J.E.

TITLE Genes differentially expressed in human prostate cancer and their use  
JOURNAL Patent: WO 0160860-A 22959 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)  
FEATURES Location/Qualifiers  
source 1..2924  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1274; DB 6; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTTAGTGTGTTATGTCCTCATTTCTATTATTAGCAATTTATTCTATGTCAT 60  
DB 480 TTTCTTTAGTGTGTTATGTCCTCATTTCTATTATTAGCAATTTATTCTATGTCAT 539  
QY 61 CCAAGAGCATTAAAGGAGTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120  
DB 540 CCAAGAGCATTAAAGGAGTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 599  
QY 121 CAGGTGACAGTCTTAATAAGTGACATGAGTGAATTTTATTTTATATATCTTTT 180  
DB 600 CAGGTGACAGTCTTAATAAGTGACATGAGTGAATTTTATTTTATATATCTTTT 659  
QY 181 TTAATCTCTATTTCT 240  
DB 660 TTAATCTCTATTTCT 719  
QY 241 TATGGCATGTAACATTTATTTATAAGTAAGTCATGTTTATATTTTCTCTGCTCT 300  
DB 720 TATGGCATGTAACATTTATTTATAAGTAAGTCATGTTTATATTTTCTCTGCTCT 779  
QY 301 CTTTATGTTATTTTTCAGAAATAGCCGTGACATGCTGACAGCAATTTACCTACCGTACC 360  
DB 780 CTTTATGTTATTTTTCAGAAATAGCCGTGACATGCTGACAGCAATTTACCTACCGTACC 839  
QY 361 TCGAAGTGTCACCATCCAGAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 840 TCGAAGTGTCACCATCCAGAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899  
QY 421 GACTTGGCGAGCTTTTTCAGAGTGTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 900 GACTTGGCGAGCTTTTTCAGAGTGTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959  
QY 481 CAATGTGACAGTGGCCATCTTGTGTTGTAGCAACTGTGCTGCTGCTGCTGCTGCTGCTGCT 540  
DB 960 CAATGTGACAGTGGCCATCTTGTGTTGTAGCAACTGTGCTGCTGCTGCTGCTGCTGCTGCT 1019  
QY 541 ACTTGGCGGGGCTTTGGATCCATTCGCAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 1020 ACTTGGCGGGGCTTTGGATCCATTCGCAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079  
QY 601 GTACTTTTCCCTGTAATATGCTCTTCTGGATGTGAATAACTCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 1080 GTACTTTTCCCTGTAATATGCTCTTCTGGATGTGAATAACTCTGCTGCTGCTGCTGCTGCTGCT 1139  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGTAGGCTTTATCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 1140 AAAGCAGACCATGAAGAGCTCTGTGAGTTTGTAGGCTTTATCTGCTGCTGCTGCTGCTGCTGCT 1199  
QY 721 TCTCTGAATATGTCAGAGCTCTCTGATGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
DB 1200 TCTCTGAATATGTCAGAGCTCTCTGATGCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1259  
QY 781 TCCATTACACCCCTACAGGAGAGGATATAGTTTTTCTTGTGCTACAGACATTAATCTTCTCT 840  
DB 1260 TCCATTACACCCCTACAGGAGAGGATATAGTTTTTCTTGTGCTACAGACATTAATCTTCTCT 1319  
QY 841 GGTGCTGTTGACGTGGTGATGATGCAAGTCTGCTGTTTGGCTTTCATCTTATGTTAGTCTTA 900

DB 1320 GGTGCTGTTGACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379  
QY 901 GAGAAACAGGAAAAATAGCATGCTCCACAGCAGATTTCTTTCGCAATCGTACAGCTGATAGA 960  
DB 1380 GAGAAACAGGAAAAATAGCATGCTCCACAGCAGATTTCTTTCGCAATCGTACAGCTGATAGA 1439  
QY 961 ACACGCAAGCAGCTGAAAAATTTTCTTACCGACTTGAGCTTAATGTTGATGAGGACGA 1020  
DB 1440 ACACGCAAGCAGCTGAAAAATTTTCTTACCGACTTGAGCTTAATGTTGATGAGGACGA 1499  
QY 1021 TTGACTTGGGAGCGACTCCTCGATCTTATTCATTAAGGAATTTGCAACAGCCATTATGAAT 1080  
DB 1500 TTGACTTGGGAGCGACTCCTCGATCTTATTCATTAAGGAATTTGCAACAGCCATTATGAAT 1559  
QY 1081 AGCGACTGTCTAGTCTTTGACACAGCATTGACAGCTTTTTCGAGAAAAATGCAATTTA 1140  
DB 1560 AGCGACTGTCTAGTCTTTGACACAGCATTGACAGCTTTTTCGAGAAAAATGCAATTTA 1619  
QY 1141 GGCATCAATGTAACTATTTCATGTTGTAATGCAATCAAAACATTTCTGCGCAGTGT 1200  
DB 1620 GGCATCAATGTAACTATTTCATGTTGTAATGCAATCAAAACATTTCTGCGCAGTGT 1679  
QY 1201 TTAAAACTTCAGTTTTCAGAGAAAAATAGGCACCCATCTGCTGCAACCTAAACCTCTTT 1260  
DB 1680 TTAAAACTTCAGTTTTCAGAGAAAAATAGGCACCCATCTGCTGCAACCTAAACCTCTTT 1739  
QY 1261 CGGTAGTGGGAAGC 1274  
DB 1740 CGGTAGTGGGAAGC 1753  
RESULT 4  
LOCUS CO493590 2924 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 25457 from Patent WO0160860.  
ACCESSION CO493590  
VERSION CO493590.1 GI:41459209  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.  
TITLE Genes differentially expressed in human prostate cancer and their use  
JOURNAL Patent: WO 0160860-A 25457 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)  
FEATURES Location/Qualifiers  
source 1..2924  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 1274; DB 6; Length 2924;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCTTTAGTGTGTTATGTCCTCATTTCTATTATTAGCAATTTATTCTATGTCAT 60  
DB 480 TTTCTTTAGTGTGTTATGTCCTCATTTCTATTATTAGCAATTTATTCTATGTCAT 539  
QY 61 CCAAGAGCATTAAAGGAGTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 120  
DB 540 CCAAGAGCATTAAAGGAGTCCACATGTTTCCGGAACATTTTGAAGAGAGCTTATC 599  
QY 121 CAGGTGACAGTCTTAATAAGTGACATGAGTGAATTTTATTTTATATATCTTTT 180  
DB 600 CAGGTGACAGTCTTAATAAGTGACATGAGTGAATTTTATTTTATATATCTTTT 659  
QY 181 TTAATCTCTATTTCT 240



```
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCTTATTCTGTCCGTCCCTGTGTCT 720
Db 1140 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGGCCTTATTCTGTCCGTCCCTGTGTCT 1199
QY 721 TCTGTAAATGGCAAGCTCTCTGGATGTGTAATGCCCCCATCTGTGATGATCAGCATAG 780
Db 1200 TCTGTAAATGGCAAGCTCTCTGGATGTGTAATGCCCCCATCTGTGATGATCAGCATAG 1259
QY 781 TCCATTACACACCTACAGGAGAGGATATAGTTTTCTTGCTACAGACATTAATCTTCT 840
Db 1260 TCCATTACACACCTACAGGAGAGGATATAGTTTTCTTGCTACAGACATTAATCTTCT 1319
QY 841 GGTGCTGTGACTGTGGTGATGATGACAGTCTGTGTTTGGCTTTTCACTTCATGTTAGTCTTA 900
Db 1320 GGTGCTGTGACTGTGGTGATGATGACAGTCTGTGTTTGGCTTTTCACTTCATGTTAGTCTTA 1379
QY 901 GAGAAACAGGAAATAACGATGGTCAACGAGCTTCTTCGCAATCGTACAGCTGATAGGA 960
Db 1380 GAGAAACAGGAAATAACGATGGTCAACGAGCTTCTTCGCAATCGTACAGCTGATAGGA 1439
QY 961 ACACGACAGCAAGCTGAAATTTTGTCTACGACTTGTGCTAAATGGTCAATAGGCGACA 1020
Db 1440 ACACGACAGCAAGCTGAAATTTTGTCTACGACTTGTGCTAAATGGTCAATAGGCGACA 1499
QY 1021 TTGACTTGGGAAGCGACTCCTCGATCTATTCTATGAAGGAATTCGCAACAGCCATTATGAAT 1080
Db 1500 TTGACTTGGGAAGCGACTCCTCGATCTATTCTATGAAGGAATTCGCAACAGCCATTATGAAT 1559
QY 1081 AGCGACTGTCTAGTCTTTCACACAGCATTTGACAGCTTTTTCGAGAAATGCAAAATGCAATTTA 1140
Db 1560 AGCGACTGTCTAGTCTTTCACACAGCATTTGACAGCTTTTTCGAGAAATGCAAAATGCAATTTA 1619
QY 1141 GGCATCAATGTAACATTTCCATGTTGTAATGGCAATCAACATTTTCGCGCAGTGT 1200
Db 1620 GGCATCAATGTAACATTTCCATGTTGTAATGGCAATCAACATTTTCGCGCAGTGT 1679
QY 1201 TTAATACTTCAGTTTCACAGAAATAAGCCACCATCTCTGCGCAACCTTAAATCTTTT 1260
Db 1680 TTAATACTTCAGTTTCACAGAAATAAGCCACCATCTCTGCGCAACCTTAAATCTTTT 1739
QY 1261 CGGTAGGTGGAAGC 1274
Db 1740 CGGTAGGTGGAAGC 1753

RESULT 6
HSA400626 31705 bp DNA linear PRI 02-APR-2001
LOCUS Homo sapiens SIAH1 gene, exons 1-2.
DEFINITION AJ400626
ACCESSION AJ400626
VERSION AJ400626.1 GI:13539602
KEYWORDS siah1 gene.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1. 31705
AUTHORS Medhioub,M., Muchardt,C., Tubacher,E., Giudicelli,C.,
Hors-Cayla,M.C. and Thomas,G.
TITLE Down regulation of the TATA-less and GC-rich SIAH1 promoter by TP53
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 31705)
AUTHORS Medhioub,M.
DIRECT SUBMISSION
TITLE Submitted (11-APR-2000) Medhioub M., Fondation Jean DAUSSET. CEPH,
JOURNAL 27, rue Juliette Dodu, Paris 75010, FRANCE
FEATURES
Location/Qualifiers
1. 31705
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="16q12-q13"
```

```
Gene 5126..30193
exon /gene="SIAH1"
5'UTR 5126..5550
evidence=experimental
number=1
intron /gene="SIAH1"
evidence=experimental
exon 5551..28322
evidence=experimental
number=2
CDS 28323..30193
evidence=experimental
number=2
3'UTR 29173..30193
/gene="SIAH1"
evidence=experimental

ORIGIN
Query Match 100.0%; Score 1274; DB 9; Length 31705;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCTTTAGTTGTTTATGTCCTCATTTTCTATTTAGCATTTTATTTCTATGTCAT 60
Db 28004 TTTCTTTAGTTGTTTATGTCCTCATTTTCTATTTAGCATTTTATTTCTATGTCAT 28063
QY 61 CCAAGACGATTAAGGGAGTTCCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 120
Db 28064 CCAAGACGATTAAGGGAGTTCCACATGTTTCCGGAACATTTTGAAAGAGAGCTTATC 28123
QY 121 CAGTGTACAGATCCCTAATAAGTGACATTCAGTGTAAATTTTATTTTATATATCTTTT 180
Db 28124 CAGTGTACAGATCCCTAATAAGTGACATTCAGTGTAAATTTTATTTTATATATCTTTT 28183
QY 181 TTAATCCTATTTTCTCTCTCTTTGCTCAGTAAATTTTGTATGAACTTTTAAAGGACT 240
Db 28184 TTAATCCTATTTTCTCTCTCTTTGCTCAGTAAATTTTGTATGAACTTTTAAAGGACT 28243
QY 241 TATGCGATGTAAACATTTTATAAGTAAGTCATGTTTATTAATTTTCTCTGCTCT 300
Db 28244 TATGCGATGTAAACATTTTATAAGTAAGTCATGTTTATTAATTTTCTCTGCTCT 28303
QY 301 CCTTATGATTTTATTTTCAGAAATGAGCGGTGACATGCTGACAGCATTTACCTACCGTACC 360
Db 28304 CCTTATGATTTTATTTTCAGAAATGAGCGGTGACATGCTGACAGCATTTACCTACCGTACC 28363
QY 361 TCGAAGTGTCCACCATCCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 28364 TCGAAGTGTCCACCATCCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 28423
QY 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTGTCTGCTTTGACTATGTGTACCGCCATCTTT 480
Db 28424 GACTTGGCGAGTCTTTTGTAGTGTCCAGTGTCTGCTTTGACTATGTGTACCGCCATCTTT 28483
QY 481 CAATGTACAGTGGCGATCTTTGTTGTAGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
```

```
Db 28484 CAATGTCAGAGTGGCCATCTTGTGTTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCA 28543
Qy 541 ACTTGC CGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATCCA 600
Db 28544 ACTTGC CGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATCCA 28603
Qy 601 GTACTTTTCCCTGTAAATATGGCTTCTGAGTGTGAATAAATCTGCGCACACAGAA 660
Db 28604 GTACTTTTCCCTGTAAATATGGCTTCTGAGTGTGAATAAATCTGCGCACACAGAA 28663
Qy 661 AAAGCAGACATCAAGCAGCTCTGTGAGTGTAGGCTTATTCTGTGCTGCGCTGGTGT 720
Db 28664 AAAGCAGACATCAAGCAGCTCTGTGAGTGTAGGCTTATTCTGTGCTGCGCTGGTGT 28723
Qy 721 TCTGTAAATGGCAAGCTCTCTGAGTGTGTAATGCCCATCTGATGATCAGCATAAG 780
Db 28724 TCTGTAAATGGCAAGCTCTCTGAGTGTGTAATGCCCATCTGATGATCAGCATAAG 28783
Qy 781 TCCATTACAACCTCAGCGAGAGGATATAGTTTTCTTGTGTACAGACATTAATCTTCT 840
Db 28784 TCCATTACAACCTCAGCGAGAGGATATAGTTTTCTTGTGTACAGACATTAATCTTCT 28843
Qy 841 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTTGGCTTTTCACTTCACTTGTAGTCTTA 900
Db 28844 GGTGCTGTTGACTGGGTGATGATGAGTCTGTTTTGGCTTTTCACTTCACTTGTAGTCTTA 28903
Qy 901 GAGAAACAGGAAATATACGATGTCACAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 960
Db 28904 GAGAAACAGGAAATATACGATGTCACAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 28963
Qy 961 ACACGCAAGCAGCTGAAAATTTTGTCTTACCGACTTGGCTAAATGGTATAGGCGACGA 1020
Db 28964 ACACGCAAGCAGCTGAAAATTTTGTCTTACCGACTTGGCTAAATGGTATAGGCGACGA 29023
Qy 1021 TTGACTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATGCAACAGCCATTATGAAT 1080
Db 29024 TTGACTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATGCAACAGCCATTATGAAT 29083
Qy 1081 AGCGACTGTCTAGTCTTTGACACACGACATTCGACAGCTTTTTCGAGAAAATGGCAATTTA 1140
Db 29084 AGCGACTGTCTAGTCTTTGACACACGACATTCGACAGCTTTTTCGAGAAAATGGCAATTTA 29143
Qy 1141 GGCATCAATGTAATTTTCCATGTTTGAATGGCAATCAAAACATTTTCTGCGCAGTGT 1200
Db 29144 GGCATCAATGTAATTTTCCATGTTTGAATGGCAATCAAAACATTTTCTGCGCAGTGT 29203
Qy 1201 TTAATACTTCAGTTTCACAGAAATAAGCACCCATCTCTGCGCAACCTTAAACTCTTT 1260
Db 29204 TTAATACTTCAGTTTCACAGAAATAAGCACCCATCTCTGCGCAACCTTAAACTCTTT 29263
Qy 1261 CGGTAGGTGGAAGC 1274
Db 29264 CGGTAGGTGGAAGC 29277

RESULT 7
AC023818 173304 bp DNA linear PRI 03-JAN-2004
LOCUS Homo sapiens chromosome 16 clone CTD-2600H12, complete sequence.
DEFINITION AC023818
ACCESSION AC023818
VERSION AC023818.5 GI:40556309
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 173304)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Unpublished
AUTHORS DOE Joint Genome Institute.
```

```
TITLE Direct Submission
JOURNAL Submitted (18-FEB-2000) Production Sequencing Facility, DOE Joint
REFERENCE Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS 3 (bases 1 to 173304)
DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-MAY-2002) Production Sequencing Facility, DOE Joint
REFERENCE Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
AUTHORS 4 (bases 1 to 173304)
DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2004) DOE Joint Genome Institute, 2800 Mitchell
COMMENT Drive, Walnut Creek, CA 94598, USA
On Jan 3, 2004 this sequence version replaced gi:20429291.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www-shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
FEATURES
Source
1..173304
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTD-2600H12"
ORIGIN
Query Match 100.0%; Score 1274; DB 9; Length 173304;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTTCTTTAGTGTCTTATGCTGCTCCATTTCTATTATTAGCATTATTATTCTATGATGCTAT 60
Db 66802 TTTCTTTAGTGTCTTATGCTGCTCCATTTCTATTATTAGCATTATTATTCTATGATGCTAT 66861
Qy 61 CCAGAGACGATTAAAGGGAGTCCACATGTTTTCGGGAACATTTTGAAGAGAGAGCTTATC 120
Db 66862 CCAGAGACGATTAAAGGGAGTCCACATGTTTTCGGGAACATTTTGAAGAGAGAGCTTATC 66921
Qy 121 CAGTGTACAGATCCCTAAATAAGTCACATTCAGTGTAAATTTTATTATTTTAAATATCTTT 180
Db 66922 CAGTGTACAGATCCCTAAATAAGTCACATTCAGTGTAAATTTTATTATTTTAAATATCTTT 66981
Qy 181 TTAATCCTATTTTCTCTCTCTTTTGTCTCAGTAAATTTTGTATGAACTTTTAAAGGACT 240
Db 66982 TTAATCCTATTTTCTCTCTCTTTTGTCTCAGTAAATTTTGTATGAACTTTTAAAGGACT 67041
Qy 241 TATGGCATGTAACACATTTATTATAAGTAAGTCATGTTTATATTTTCTCTGCTCT 300
Db 67042 TATGGCATGTAACACATTTATTATAAGTAAGTCATGTTTATATTTTCTCTGCTCT 67101
Qy 301 CCTTATGATTTTATTCAGAAATAGCCGTACAGCTGTACAGCAATTA CTAACCGGTACC 360
Db 67102 CCTTATGATTTTATTCAGAAATAGCCGTACAGCTGTACAGCAATTA CTAACCGGTACC 67161
Qy 361 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCGCTGCTGCGCACACTGCGATCCCAACAT 420
Db 67162 TCGAAGTGTCCACCATCCAGAGGGTGCCTGCGCTGCTGCGCACACTGCGATCCCAACAT 67221
Qy 421 GACTTGGCGAGTCTTTTGTGAGTGTCCAGTCTGCTTGTGACTATGTTGTACCGCCATCTTT 480
Db 67222 GACTTGGCGAGTCTTTTGTGAGTGTCCAGTCTGCTTGTGACTATGTTGTACCGCCATCTTT 67281
Qy 481 CAATGTACAGAGTGGCCATCTCTGTTTGTAGCAACTGTGCGCCCAAGCTCA CATGTTGTCCA 540
Db 67282 CAATGTACAGAGTGGCCATCTCTGTTTGTAGCAACTGTGCGCCCAAGCTCA CATGTTGTCCA 67341
Qy 541 ACTTGC CGGGGCCCTTTGGGATCCATTCGCAACTTGGCTATGAGAAAGTGGCTAATCCA 600
```



Db 67342 ACTTGCCTGGGCGCCCTTGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAATTCA 67401  
QY 601 GTACTTTTCCCTGTAAATATATCGCTCTCTGATGTGAAATAAATCTGCCACACAGAA 660  
Db 67402 GTACTTTTCCCTGTAAATATATCGCTCTCTGATGTGAAATAAATCTGCCACACAGAA 67461  
QY 661 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGCCCTTATTCCTGCTCCGTCCTGGTGT 720  
Db 67462 AAAGCAGACCATGAAGAGCTCTGTGAGTTTAGCCCTTATTCCTGCTCCGTCCTGGTGT 67521  
QY 721 TCCTGTAATGCAAGGCTCTCTGATGTGTAATGTCCTGATGCAAGCATGAGCATGAGTAAG 780  
Db 67522 TCCTGTAATGCAAGGCTCTCTGATGTGTAATGTCCTGATGCAAGCATGAGCATGAGTAAG 67581  
QY 781 TCCTGTAATGCAAGGCTCTCTGATGTGTAATGTCCTGATGCAAGCATGAGCATGAGTAAG 840  
Db 67582 TCCTGTAATGCAAGGCTCTCTGATGTGTAATGTCCTGATGCAAGCATGAGCATGAGTAAG 67641  
QY 841 GGTGCTGTTGACTGGGTGATGATGAGTCCCTGCTTTTGGCTTTTCACTTTCATGTTAGTCTTA 900  
Db 67642 GGTGCTGTTGACTGGGTGATGATGAGTCCCTGCTTTTGGCTTTTCACTTTCATGTTAGTCTTA 67701  
QY 901 GAGAAACAGGAAATAATGATGCTGACAGGAGTCTTTCGCAATCGTACAGCTGATAGGA 960  
Db 67702 GAGAAACAGGAAATAATGATGCTGACAGGAGTCTTTCGCAATCGTACAGCTGATAGGA 67761  
QY 961 ACACGACAGGAGCTGAAATTTTGTACCGACTGAGTAAATGGTCAATAGGCGAGGA 1020  
Db 67762 ACACGACAGGAGCTGAAATTTTGTACCGACTGAGTAAATGGTCAATAGGCGAGGA 67821  
QY 1021 TTGACTTGGGAAGGAGTCTCTCGATCTATTCTATGAGGAATTTGCAACAGCCATTTAAT 1080  
Db 67822 TTGACTTGGGAAGGAGTCTCTCGATCTATTCTATGAGGAATTTGCAACAGCCATTTAAT 67881  
QY 1081 AGCGACTGTCTAGTCTTTCGACACAGCAATGACAGTCTTTTGGCAAAATGGCAATTTA 1140  
Db 67882 AGCGACTGTCTAGTCTTTCGACACAGCAATGACAGTCTTTTGGCAAAATGGCAATTTA 67941  
QY 1141 GGCATCAATGTAACATTTTCCATGTTGAAATGGCAATTTTGGCAAAATGGCAATTTA 1200  
Db 67942 GGCATCAATGTAACATTTTCCATGTTGAAATGGCAATTTTGGCAAAATGGCAATTTA 68001  
QY 1201 TTAACACTTTCAGTTTTCACAGAAATAAGGACCCATCTGCTGCCAACCTTAAACTCTTT 1260  
Db 68002 TTAACACTTTCAGTTTTCACAGAAATAAGGACCCATCTGCTGCCAACCTTAAACTCTTT 68061  
QY 1261 CGTAGGTGGAAGC 1274  
Db 68062 CGTAGGTGGAAGC 68075

RESULT 8  
AX833145  
LOCUS 2829 bp DNA linear PAT 15-DEC-2003  
DEFINITION Sequence 269 from Patent EP1347046.  
ACCESSION AX833145  
VERSION AX833145.1 GI:39919280  
KEYWORDS Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato H., Ishii,S., Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and Masuho,Y.  
Full-length cDNA sequences  
TITLE Patent: EP 1347046-A 269 24-SEP-2003;  
JOURNAL Research Association for Biotechnology (JP)  
FEATURES  
Location/Qualifiers  
1..2829  
/organism="Homo sapiens"

/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 96.0%; Score 1223; DB 6; Length 2829;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTTCTTTAGTTGTTTATGGTCCATTTTCTATTATTAGCAATTTATTCTATGTAAGTCTAT 60  
Db 1134 TTTCTTTAGTTGTTTATGGTCCATTTTCTATTATTAGCAATTTATTCTATGTAAGTCTAT 1193  
QY 61 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAACATTTTGAAGAGAGCTTATC 120  
Db 1194 CCAAGACGATTAAGGGAGTTCCACATGTTTTCGGAACATTTTGAAGAGAGCTTATC 1253  
QY 121 CAGTGTACAGATCCTTAATAAAGTGACATTCAGTGTAAATTTTATTATTATATCTTTT 180  
Db 1254 CAGTGTACAGATCCTTAATAAAGTGACATTCAGTGTAAATTTTATTATTATATCTTTT 1313  
QY 181 TTAATCCTATTCTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAA CTTTAAAAAGGACT 240  
Db 1314 TTAATCCTATTCTTCTCTCTTTTGTCTCAGTAAATTTTGTATGAAA CTTTAAAAAGGACT 1373  
QY 241 TATGCATGTAAACATTTATATAAGTAAGTCAATGTTTATTTTCTCTGCT 300  
Db 1374 TATGCATGTAAACATTTATATAAGTAAGTCAATGTTTATTTTCTCTGCT 1433  
QY 301 CCTATGATTTTATTTAGAAATGAGCGTACAGCTGCTACAGCAATACCTACCGGTACC 360  
Db 1434 CCTATGATTTTATTTTAGAAATGAGCGTACAGCTGCTACAGCAATACCTACCGGTACC 1493  
QY 361 TCGAAGTGTCCACCATCCAGAGGGTGTCTGCTTGTAGCACTGGGCACCACTGCATCCAACT 420  
Db 1494 TCGAAGTGTCCACCATCCAGAGGGTGTCTGCTTGTAGCACTGGGCACCACTGCATCCAACT 1553  
QY 421 GACTTGGCGAGTCTTTTGTAGTGTCCAGTGTCTTGTGATATGTTTACCGCCATCTTT 480  
Db 1554 GACTTGGCGAGTCTTTTGTAGTGTCCAGTGTCTTGTGATATGTTTACCGCCATCTTT 1613  
QY 481 CAATGTACAGTGGCGCATCTTTTGTAGCACTGTCCGCCAAAGCTACATGTTGTCCA 540  
Db 1614 CAATGTACAGTGGCGCATCTTTTGTAGCACTGTCCGCCAAAGCTACATGTTGTCCA 1673  
QY 541 ACTTCCCGGGGCTTTGGGATCCATTCGCACTTTGGCTATGGAGAAAGTGGCTAATTCA 600  
Db 1674 ACTTCCCGGGGCTTTGGGATCCATTCGCACTTTGGCTATGGAGAAAGTGGCTAATTCA 1733  
QY 601 GTACTTTTCCCTGTAAATATATGCGTCTTCTGGATGTGAAATAAATCTTGCCACACAGAA 660  
Db 1734 GTACTTTTCCCTGTAAATATATGCGTCTTCTGGATGTGAAATAAATCTTGCCACACAGAA 1793  
QY 661 AAAGCAGACCATGAAGAGCTCTGTCAGTTTAGCCCTTATTCCTGCTCCGTCCTGGTGTCT 720  
Db 1794 AAAGCAGACCATGAAGAGCTCTGTCAGTTTAGCCCTTATTCCTGCTCCGTCCTGGTGTCT 1853  
QY 721 TCCTGTAATGCAAGGCTCTCTGATGTGTAATGTCCTGATGCAAGCATGAGCATGAGTAAG 780  
Db 1854 TCCTGTAATGCAAGGCTCTCTGATGTGTAATGTCCTGATGCAAGCATGAGCATGAGTAAG 1913  
QY 781 TCCTGTAATGCAAGGCTCTCTGATGTGTAATGTCCTGATGCAAGCATGAGCATGAGTAAG 840  
Db 1914 TCCTGTAATGCAAGGCTCTCTGATGTGTAATGTCCTGATGCAAGCATGAGCATGAGTAAG 1973  
QY 841 GGTGCTGTTGACTGGGTGATGATGAGTCCCTGCTTTTGGCTTTTCACTTTCATGTTAGTCTTA 900  
Db 1974 GGTGCTGTTGACTGGGTGATGATGAGTCCCTGCTTTTGGCTTTTCACTTTCATGTTAGTCTTA 2033  
QY 901 GAGAAACAGGAAATAATGATGCTGACAGGAGTCTTTCGCAATCGTACAGCTGATAGGA 960  
Db 2034 GAGAAACAGGAAATAATGATGCTGACAGGAGTCTTTCGCAATCGTACAGCTGATAGGA 2093  
QY 961 ACACGACAGGAGCTGAAATTTTGTACCGACTGAGTAAATGGTCAATAGGCGAGGA 1020



Db 2094 ACACGGACGACCTGAAATTTTGGCTTCCCGACTTGGAGCTTAATGGTCATAGGCGACGA 2153  
Qy 1021 TTGACTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATTGCAACAGCCATTATGAAT 1080  
Db 2154 TTGACTTGGGAAGCGACTCTCGATCTATTTCATGAAGGAATTGCAACAGCCATTATGAAT 2213  
Qy 1081 AGGACTGTCTAGCTTTTGACACAGCATTTGACAGCTTTTTCAGAGAAATGGCAATTTA 1140  
Db 2214 AGGACTGTCTAGCTTTTGACACAGCATTTGACAGCTTTTTCAGAGAAATGGCAATTTA 2273  
Qy 1141 GGCAATCAATCTACTATTTCATGTGTGAAATGGCAATCAACATTTTTCGCCAGTGT 1200  
Db 2274 GGCAATCAATCTACTATTTCATGTGTGAAATGGCAATCAACATTTTTCGCCAGTGT 2333  
Qy 1201 TTAATAACTTCAGTTTTCACAGAAATAAGGCACCCATCTGTCTGCCAACCTTAAACTCTTTT 1260  
Db 2334 TTAATAACTTCAGTTTTCACAGAAATAAGGCACCCATCTGTCTGCCAACCTTAAACTCTTTT 2393  
Qy 1261 CGGTAGGTGAAGC 1274  
Db 2394 CGGTAGGTGAAGC 2407

## RESULT 9

AK094663  
LOCUS  
DEFINITION Homo sapiens CDNA FLJ37344 fis, clone BRAMY2021139, highly similar to Seven in absentia.  
ACCESSION AK094663  
VERSION AK094663.1 GI:21753769  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohyashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoi, R., Kaku, Y., Kodaira, H., Kondo, H., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hiro, K., Ohmori, Y., Kawabata, A., Hiki, J., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs  
Nat. Genet. 36 (1), 40-45 (2004)

## TITLE

JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,

Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamuro, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.  
NEDO human cDNA sequencing project  
Unpublished

TITLE  
JOURNAL  
REFERENCE

3 (bases 1 to 2829)

## AUTHORS

Isogai, T. and Yamamoto, J.

## JOURNAL

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7

Kazusa-genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986

(E-mail: genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; CDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: HRI and RAB.

## FEATURES

Location/Qualifiers

1..2829

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="BRAMY2021139"

/tissue\_type="amygdala"

/clone\_lib="BRAMY2"

/note="cloning vector: pME18SFL3"

## ORIGIN

Query Match 96.0%; Score 1223; DB 9; Length 2829;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1273; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTTCTTTAGTGTGTTATGGTCCATTTTCTATTTTATTTAGCATTTTATTTCTATGTAGTCTAT 60

Db 1134 TTTCTTTAGTGTGTTATGGTCCATTTTCTATTTTATTTAGCATTTTATTTCTATGTAGTCTAT 1193

Qy 61 CCAAGACGATTAAGGGAGTCCACATGTTTTCGGAAACATTTTGAAGAGAGAGCTTATC 120

Db 1194 CCAAGACGATTAAGGGAGTCCACATGTTTTCGGAAACATTTTGAAGAGAGAGCTTATC 1253

Qy 121 CAGGTACAGATCCTAATAAGTGACATTCAGTGAATTTTATTTTATATATCTTTT 180

Db 1254 CAGGTACAGATCCTAATAAGTGACATTCAGTGAATTTTATTTTATATATCTTTT 1313

Qy 181 TTAATCCTATTTTCTCTCTTTTCTCAGTAAATTTTGTATGAACCTTTAAAGGACT 240

Db 1314 TTAATCCTATTTTCTCTCTTTTCTCAGTAAATTTTGTATGAACCTTTAAAGGACT 1373

Qy 241 TATGGCATGTAACCAATTTATTAAGTAAGTCATGGTTATATTTTCTCTCGCT 300

Db 1374 TATGGCATGTAACCAATTTATTAAGTAAGTCATGGTTATATTTTCTCTCGCT 1433

Qy 301 CCTTATGATTTTATTTTCAGAAATAGCCGTCAGATGCTACAGCATTTACCTACCGTACC 360

Db 1434 CCTTATGATTTTATTTTCAGAAATAGCCGTCAGATGCTACAGCATTTACCTACCGTACC 1493

Qy 361 TCGAAGTGTCCACCATCCAGAGGGTCCCTGCCCTGACCTGGCAGCACTGCATCCACAAT 420

Db 1494 TCGAAGTGTCCACCATCCAGAGGGTCCCTGCCCTGACCTGGCAGCACTGCATCCACAAT 1553

Qy 421 GACTTGGGCGAGTCTTTTTCAGTGTCCAGTCTCTTTTGAATGTGTGTACGGCCCATCTTT 480

Db 1554 GACTTGGGCGAGTCTTTTTCAGTGTCCAGTCTCTTTTGAATGTGTGTACGGCCCATCTTT 1613

Qy 481 CAATGTACAGAGTGGCCATCTTTTGTAGCAACTGTGCCCCAAAGCTCATGTGTGTCCA 540

Db 1614 CAATGTACAGAGTGGCCATCTTTTGTAGCAACTGTGCCCCAAAGCTCATGTGTGTCCA 1673

```
QY 541 ACTTGGCGGGCCCTTTGGGATCCATTGCAACTTTGGCTATGAGAAAGTGGCTAATTCA 600
Db 1674 ACTTGGCGGGCCCTTTGGGATCCATTGCAACTTTGGCTATGAGAAAGTGGCTAATTCA 1733
QY 601 GTACTTTTCCCTCTGTAATATATCGTCTTCTGATGTGAAATAACTCTGCACACACAGAA 660
Db 1734 GTACTTTTCCCTCTGTAATATATCGTCTTCTGATGTGAAATAACTCTGCACACACAGAA 1793
QY 661 AAGCAGACCATGAGAGCTCTGTGAGTTAGGCTTATTCCTGTCCGTCGGCTGGTCT 720
Db 1794 AAGCAGACCATGAGAGCTCTGTGAGTTAGGCTTATTCCTGTCCGTCGGCTGGTCT 1853
QY 721 TCCCTGTAATGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATGATCAGCATTAAG 780
Db 1854 TCCCTGTAATGCAAGGCTCTCTGATGCTGTAATGCCCATCTGATGATCAGCATTAAG 1913
QY 781 TCCATTACAACCTACAGGAGAGGATATAGTTTTTCTTGCTACAGACATTAATCTTCT 840
Db 1914 TCCATTACAACCTACAGGAGAGGATATAGTTTTTCTTGCTACAGACATTAATCTTCT 1973
QY 841 GGTGCTGTGACTGGTGTGATGATGATGCTCTGTTTGGCTTTTCACTTCATGTTAGTCTTA 900
Db 1974 GGTGCTGTGACTGGTGTGATGATGATGCTCTGTTTGGCTTTTCACTTCATGTTAGTCTTA 2033
QY 901 GAGAAACAGGAAATAACGATGCTCACAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 960
Db 2034 GAGAAACAGGAAATAACGATGCTCACAGCAGTCTTTCGCAATCGTACAGCTGATAGGA 2093
QY 961 ACACGCAAGCAAGCTGAAATTTTCTTACCGACTTGAGCTAAATGGTCATAGGCGAGGA 1020
Db 2094 ACACGCAAGCAAGCTGAAATTTTCTTACCGACTTGAGCTAAATGGTCATAGGCGAGGA 2153
QY 1021 TTGACTTGGGAGCGACTCTCGATCTATTTCATGAGGAATTCACACGCCATTAATGAAT 1080
Db 2154 TTGACTTGGGAGCGACTCTCGATCTATTTCATGAGGAATTCACACGCCATTAATGAAT 2213
QY 1081 AGCGACTGTCTAGTCTTTCACACCGACCATTCACACAGCTTTTTCGAGAAATTCGCAATTTA 1140
Db 2214 AGCGACTGTCTAGTCTTTCGACCGACCATTCACACAGCTTTTTCGAGAAATTCGCAATTTA 2273
QY 1141 GGCATCAATGTAACTATTTCATGTTGGAATGGAATCAAACTTTTCGCGCAGTGT 1200
Db 2274 GGCATCAATGTAACTATTTCATGTTGGAATGGAATCAAACTTTTCGCGCAGTGT 2333
QY 1201 TTAATACTTCAGTTTCACAGAAATAAGCCACCTCTGCTGCAACCTTAACTCTTT 1260
Db 2334 TTAATACTTCAGTTTCACAGAAATAAGCCACCTCTGCTGCAACCTTAACTCTTT 2393
QY 1261 CGGTAGTGGGAAGC 1274
Db 2394 CGGTAGTGGGAAGC 2407

RESULT 10
LOCUS HSM807215
DEFINITION Homo sapiens mRNA; cDNA DKFZp686L1897 (from clone DKFZp686L1897);
complete cds.
ACCESSION BX647064
VERSION BX647064.1 GI:34366097
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2972)
AUTHORS Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Oeinger,A.,
Fobo,G., Han,M. and Wiemann,S.
CONSTRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuerberg, GERMANY
```

```
COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp686L1897) is available at the RZPD in Berlin. Please contact
the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1..2972
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="16q12.1"
/clone="DKFZp686L1897"
/tissue_type="human retina"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
1..2972
/gene="DKFZp686L1897"
/feature="1910
/gene="DKFZp686L1897"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE46191.1"
/db_xref="GI:34366098"
/translation="MSRQTATATLPTGTSKCPSPQRPVLPALTTASNNDLASLFECPVC
FDYVLPILQOSGHLVCSNCRPLTCCPTCRPLGSIRNLAMKAVNSVLPFPKYAS
SGCEITLPHTEKHEELCEFRPYSCPCGASCKWQSLDVAVPMHMQHKSITTLQ
EDIVFLATDINLSGAVDWMQSCFGPHFMLEVLEKQEKYDGHQFPFAIVOLIGTKQA
ENFAYRLNGLHRRRLTWEATPSRHGGIATAIMNSDCLVFTSIAQLFAENGNLGIN
VTISM"
2931..2936
/polyA_signal
/gene="DKFZp686L1897"
2948
/polyA_site
/gene="DKFZp686L1897"
ORIGIN
Query Match 92.0%; Score 1172; DB 9; Length 2972;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTTCTTTAGTTGTTTATGTCCTCATTTTCTATTTAGCATTTATTTATTTAGTCTAT 60
Db 741 TTTCTTTAGTTGTTTATGTCCTCATTTTCTATTTAGCATTTATTTATTTAGTCTAT 800
QY 61 CCAAGACGATTAAAGGAGTCCACATGTTTTCGGAAACATTTTGAAGAGAGCTTATC 120
Db 801 CCAAGACGATTAAAGGAGTCCACATGTTTTCGGAAACATTTTGAAGAGAGCTTATC 860
QY 121 CAGTGTACGATCCTTAATAAGTGACATTCAGTGTAAATTTATTTTATATCTTTT 180
Db 861 CAGTGTACGATCCTTAATAAGTGACATTCAGTGTAAATTTATTTTATATCTTTT 920
QY 181 TTAATCCTATTTTCTCTCTCTTTTCTCAGTAAATTTTGTATGAACTTTAAAGGACT 240
Db 921 TTAATCCTATTTTCTCTCTCTTTTCTCAGTAAATTTTGTATGAACTTTAAAGGACT 980
QY 241 TATGGCATGTAAACATTTTATAAAGTAAGTCATGGTTATTAATTTTCTCTGCTCT 300
Db 981 TATGGCATGTAAACATTTTATAAAGTAAGTCATGGTTATTAATTTTCTCTGCTCT 1040
QY 301 CTTTATGATTTTATTTTTCAGAAATGCGCTCAGCTCAGCTACAGCATTTACCTCCGCTACC 360
Db 1041 CTTTATGATTTTATTTTTCAGAAATGCGCTCAGCTCAGCTACAGCATTTACCTCCGCTACC 1100
QY 361 TCGAAGTGTCCACCATCCAGAGGGTCTCTGCGCTGACTGGCAGCAACTGCATCCCAACAT 420
Db 1101 TCGAAGTGTCCACCATCCAGAGGGTCTCTGCGCTGACTGGCAGCAACTGCATCCCAACAT 1160
```

QY	421	GACTTGGCGAGCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTGTGTACGGCCCAATCTTT	480
Db	1161	GACTTGGCGAGCTTTTTCAGTGTCCAGTCTGCTTTGACTATGTGTGTACGGCCCAATCTTT	1220
QY	481	CAATGTCCAGTGGCCATCTTTTGTAGCAACTGTGCGCCCAAGCTCACAATGTTCTTCCA	540
Db	1221	CAATGTCCAGTGGCCATCTTTTGTAGCAACTGTGCGCCCAAGCTCACAATGTTCTTCCA	1280
QY	541	ACTTGGCGGGGCCCTTTGGATCCATTCGCAACTTCGGCTATGAGAAAGTGGCTAATCCA	600
Db	1281	ACTTGGCGGGGCCCTTTGGATCCATTCGCAACTTCGGCTATGAGAAAGTGGCTAATCCA	1340
QY	601	GTACTTTTCCCTGTAATAATATGGCTTCTTGGATGGAATAACTTCGCCACACACAGAA	660
Db	1341	GTACTTTTCCCTGTAATAATATGGCTTCTTGGATGGAATAACTTCGCCACACACAGAA	1400
QY	661	AAAGCAGACATCAAGAGCTCTGTGAGTTTATGAGCCCTTATTCCTGTCGCTGCCCTGGTGT	720
Db	1401	AAAGCAGACATCAAGAGCTCTGTGAGTTTATGAGCCCTTATTCCTGTCGCTGCCCTGGTGT	1460
QY	721	TCTGTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGTATGCATCAGCATAAG	780
Db	1461	TCTGTAAATGGCAAGGCTCTCTGGATGCTGTAAATGCCCATCTGTATGCATCAGCATAAG	1520
QY	781	TCCATTACAAACCTCAGGAGAGGATATAGTTTTTCTTGTCTCAGACATTAATCTTCTCT	840
Db	1521	TCCATTACAAACCTCAGGAGAGGATATAGTTTTTCTTGTCTCAGACATTAATCTTCTCT	1580
QY	841	GGTGTCTGTGACTGGGTGATGATGAGTCTGCTGTTTGGCTTTTCACTTCATGTTAGTCTTA	900
Db	1581	GGTGTCTGTGACTGGGTGATGATGAGTCTGCTGTTTGGCTTTTCACTTCATGTTAGTCTTA	1640
QY	901	GAGAAACAGGAAATAACGATGGTCCACGAGCTTCTTCGCAATCTGTACAGCTGTATGAGA	960
Db	1641	GAGAAACAGGAAATAACGATGGTCCACGAGCTTCTTCGCAATCTGTACAGCTGTATGAGA	1700
QY	961	ACACGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGTCATAGGAGACGA	1020
Db	1701	ACACGCAAGCAAGCTGAAATTTTGTCTTACCGACTTGAGCTAAATGGTCATAGGAGACGA	1760
QY	1021	TTGACTTGGGAAGCACTCTCGATCTATTCATGAGGAATGCAACAGCCATATGAT	1080
Db	1761	TTGACTTGGGAAGCACTCTCGATCTATTCATGAGGAATGCAACAGCCATATGAT	1820
QY	1081	AGCGACTGTCTAGTCTTTGCACACCACTTGCACAGCTTTTTCGCAAAATGCAATTTA	1140
Db	1821	AGCGACTGTCTAGTCTTTGCACACCACTTGCACAGCTTTTTCGCAAAATGCAATTTA	1880
QY	1141	GGCATCAATGTAACTATTTCCATGTTGAAATGGCAATCAAACTTTTCGCCAGTGT	1200
Db	1881	GGCATCAATGTAACTATTTCCATGTTGAAATGGCAATCAAACTTTTCGCCAGTGT	1940
QY	1201	TTAAATCTTCAAGTTTCAAGAAATTAAGGACCCATCTGTCTGCCAATCTAAATCTTTT	1260
Db	1941	TTAAATCTTCAAGTTTCAAGAAATTAAGGACCCATCTGTCTGCCAATCTAAATCTTTT	2000
QY	1261	CGTAGTGGAGC 1274	
Db	2001	CGTAGTGGAGC 2014	
RESULT	11		
LOCUS	AK056051/c		
DEFINITION	Homo sapiens cDNA FLJ31489 fis, clone NT2NE2003308.	2128 bp	mRNA linear PRI 30-JAN-2004
ACCESSION	AK056051		
VERSION	AK056051.1	GI:16551141	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		

## AUTHORS

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yanashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project

Unpublished

3 (bases 1 to 2128)

Isogai, T., Otsuki, T. and Sugiyama, T.

Direct Submission

Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES  
source

1. .2128  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2NE2003308"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2NE2"  
/note="cloning vector: pME18FL3-mRNA from NT2 neuron after the differentiation of NT2 neuronal precursor cells."

## ORIGIN



834 TCTTCTGCTGCTTTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTTCACATTCATGTT 893  
Db TCTTCTGCTGCTTTGACTGGGTGATGATGAGTCTCTGTTTGGCTTTTCACATTCATGTT 1126  
894 AGTCTTAGAGAAACAGAGAAAATACGATGCTCACCAGCAGTCTTTCGCAATCGTACAGCT 953  
Db AGTCTTAGAGAAACAGAGAAAATACGATGCTCACCAGCAGTCTTTCGCAATCGTACAGCT 1186  
954 GATAGGAACACGACGACGCTGAAATTTGCTTACCGACTTGAGCTAAATGGTTCATAG 1013  
Db GATAGGAACACGACGACGCTGAAATTTGCTTACCGACTTGAGCTAAATGGTTCATAG 1246  
1014 GCGACGATTCAGCTTGGGAGCGACTCTCGATCTATTCATGAAGGATTCGACACGCCAT 1073  
Db GCGACGATTCAGCTTGGGAGCGACTCTCGATCTATTCATGAAGGATTCGACACGCCAT 1306  
1074 TATGAATAGGAGCTGCTAGTCTTTGACACACGACATTCGACAGCTTTTTCGAGAAATGG 1133  
Db TATGAATAGGAGCTGCTAGTCTTTGACACACGACATTCGACAGCTTTTTCGAGAAATGG 1366  
1134 CAATTTAGGCATCAATGTAATCTTTCCATGCTTGAATGCGCAATCAACATTTTCTCG 1193  
Db CAATTTAGGCATCAATGTAATCTTTCCATGCTTGAATGCGCAATCAACATTTTCTCG 1426  
1194 CAGTGTTTAAACTTCAGTTTTCACAGAAATTAAGCAGCCATCTCTGTCGCAACCTAAA 1253  
Db CAGTGTTTAAACTTCAGTTTTCACAGAAATTAAGCAGCCATCTCTGTCGCAACCTAAA 1486  
1254 ACTCTTTTCGGTAGTGGAGC 1274  
Db ACTCTTTTCGGTAGTGGAGC 1507

RESULT 13  
BC035562  
LOCUS  
DEFINITION  
Homo sapiens seven in absentia homolog 1 (Drosophila), mRNA (cDNA  
clone MGC:40343 IMAGE:5224755), complete cds.  
ACCESSION  
BC035562  
VERSION  
BC035562.1 GI:23274141  
KEYWORDS  
MGC.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1540)  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,  
Schetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S.,  
Carninci P., Prange C., Raja S.S., Loquellano N.A., Peters G.J.,  
Abramson R.D., Mullah S.J., Bosak S.A., McEwan P.J.,  
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,  
Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S.,  
Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,  
Boutard G.G., Blakesley R.W., Touchman J.W., Green E.D.,  
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S., Krzywinski M.I., Skalska U., Smalls D.E.,  
Scherch A., Schein J.E., Jones S.J. and Marra M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 1540)  
Strausberg R.  
Direct Submission  
Submitted (31-JUL-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)  
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Place: 64 Row: 0 Column: 8  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4506946.

## FEATURES

Location/Qualifiers  
1..1540  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="NIH MGC\_120"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..1540  
/gene="SIAH1"  
/db\_xref="LocusID:6477"  
/db\_xref="MIM:602212"  
118..966  
/gene="SIAH1"  
/codon\_start=1  
/product="SIAH1 protein"  
/protein\_id="AAH35562.1"  
/db\_xref="GI:23274142"  
/db\_xref="LocusID:6477"  
/db\_xref="MIM:602212"  
misc\_feature  
361..963  
/gene="SIAH1"  
/note="Sina: Region: Seven in absentia protein family. The  
seven in absentia (sina) gene was first identified in  
Drosophila. The Drosophila Sina protein is essential for  
the determination of the R7 pathway in photoreceptor cell  
development: the loss of functional Sina results in the  
transformation of the R7 precursor cell to a non-neuronal  
cell type. The Sina protein contains an N-terminal RING  
finger domain pfam00097. Through this domain, Sina binds  
E2 ubiquitin-conjugating enzymes (UbcD1). Sina also  
interacts with Tramtrack (TTK88) via PHYL. Tramtrack is a  
transcriptional repressor that blocks photoreceptor  
determination, while PHYL down-regulates the activity of  
TTK88. In turn, the activity of PHYL requires the  
activation of the Sevenless receptor tyrosine kinase, a  
process essential for R7 determination. It is thought that  
thus Sina targets TTK88 for degradation, therefore





```
QY 737 GCTCTCTGGATGCTGTAAATGCCCATCTGATGATCATCAGCATTAAGTCCATTACCAACCTTAC 796
Db 421 GCTCTCTGGATGCTGTAAATGCCCATCTGATGATCATCAGCATTAAGTCCATTACCAACCTTAC 480
QY 797 AGGAGAGGAGATAGTATTTTCTTGCTACAGACATTAATCTTCTCTGGTGTCTGTGACTGGG 856
Db 481 AGGAGAGGAGATAGTATTTTCTTGCTACAGACATTAATCTTCTCTGGTGTCTGTGACTGGG 540
QY 857 TGATGATGAGTCTGTTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAAAT 916
Db 541 TGATGATGAGTCTGTTTGGCTTTTCACTTCATGTTAGTCTTAGAGAAACAGGAAAAAT 600
QY 917 ACCATGTCACACGAGCTTCTCGCAATCTCAGCTGATAGGAAACAGCAAGCTG 976
Db 601 ACCATGTCACACGAGCTTCTCGCAATCTCAGCTGATAGGAAACAGCAAGCTG 660
QY 977 AAAATTTTGTCTTACCGACTTGAAGTAAATGGTCAATAGGCGACGATTTGCGAAGCGA 1036
Db 661 AAAATTTTGTCTTACCGACTTGAAGTAAATGGTCAATAGGCGACGATTTGCGAAGCGA 720
QY 1037 CTCTCTGATCTATTCATGAAGAAATGCAACAGCCATTATGAATAGCGACTGTCTAGTCT 1096
Db 721 CTCTCTGATCTATTCATGAAGAAATGCAACAGCCATTATGAATAGCGACTGTCTAGTCT 780
QY 1097 TTGACACCGAGCTGACAGCTTTTTCGCAAAATGCAATTTAGGCATCAATGTAACCTA 1156
Db 781 TTGACACCGAGCTGACAGCTTTTTCGCAAAATGCAATTTAGGCATCAATGTAACCTA 840
QY 1157 TTTTCCATGTGTGAATGGCAATCAAACTTTTCTGCGCAGTGTTTAAACTTTTCACTTTC 1216
Db 841 TTTTCCATGTGTGAATGGCAATCAAACTTTTCTGCGCAGTGTTTAAACTTTTCACTTTC 900
QY 1217 ACAGAAATAAGGACCCCATCTGCTGCTGCCAACCTTAAACTTTTTCGGTAGGTGGAAGC 1274
Db 901 ACAGAAATAAGGACCCCATCTGCTGCTGCCAACCTTAAACTTTTTCGGTAGGTGGAAGC 958

RESULT 15
HSU76247 2048 bp mRNA linear PRI 10-DEC-1997
LOCUS Human hSIAH1 mRNA, complete cds.
DEFINITION
ACCESSION U76247
VERSION U76247.1 GI:2673965
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hu, G., Chung, Y.-L., Glover, T., Valentine, V., Look, A.T. and
Fearon, E.R.
Characterization of human homologs of the Drosophila seven in
absentia (sina) gene
Genomics 46 (1), 103-111 (1997)
98066768
9403064
2 (bases 1 to 2048)
Hu, G., Chung, Y.-L., Look, A.T. and Fearon, E.R..
Direct Submission
Submitted (24-OCT-1996) Internal Medicine, University of Michigan,
4301 MSRB III, BOX 0638, 1150 W. Medical Center Dr., Ann Arbor, MI
48109-0638, USA
FEATURES
Location/Qualifiers
1..2048
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q12"
187..1035
/note="sina homologue 1"
/codon_start=1
/product="hSIAH1"

CDS
```

```
ORIGIN
Query Match 75.2%; Score 958; DB 9; Length 2048;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 958; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 CAGAAATGAGCCGTACAGACTGTCTACAGCAATTACCTACCGGTACCTCGAAGTGTCCACCAT 376
Db 182 CAGAAATGAGCCGTACAGACTGTCTACAGCAATTACCTACCGGTACCTCGAAGTGTCCACCAT 241
QY 377 CCAGAGGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
Db 242 CCAGAGGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
QY 437 TTGAGTGTCCAGTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
Db 302 TTGAGTGTCCAGTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
QY 497 ATCTTGTGTTGAGCAACTGTGCGCCAAAGCTCACTGTTGTGCCAATCTGCGCGGGGCCCTT 556
Db 362 ATCTTGTGTTGAGCAACTGTGCGCCAAAGCTCACTGTTGTGCCAATCTGCGCGGGGCCCTT 421
QY 557 TGGGATCCATTTGCAACTTGGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCCTGTA 616
Db 422 TGGGATCCATTTGCAACTTGGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCCTGTA 481
QY 617 AATATCGGTCTTCTGCAATGTGAAATACTCTGCCACACACAGAAAGAGAGACCATGAAG 676
Db 482 AATATCGGTCTTCTGCAATGTGAAATACTCTGCCACACACAGAAAGAGAGACCATGAAG 541
QY 677 AGCTCTGTGAGTTTAGCCCTTATTCCTGTCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
Db 542 AGCTCTGTGAGTTTAGCCCTTATTCCTGTCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
QY 737 GCTCTCTGATGCTGTAAATGCCCCCATCTGATGATCAGCATTAAGTCCATTACCAACCTTAC 796
Db 602 GCTCTCTGATGCTGTAAATGCCCCCATCTGATGATCAGCATTAAGTCCATTACCAACCTTAC 661
QY 797 AGGAGAGATATAGTATTTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCT 856
Db 662 AGGAGAGATATAGTATTTTCTGCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCTGCT 721
QY 857 TGATGATGAGTCTGCTGCTTGTGCTTTCATCTTCATGTTAGTCTTAGAGAAACAGGAAAAAT 916
Db 722 TGATGATGAGTCTGCTGCTTGTGCTTTCATCTTCATGTTAGTCTTAGAGAAACAGGAAAAAT 781
QY 917 ACAGTGGTCAACGAGCAGTCTTTCGCAATCTGATGATGATGATGATGATGATGATGATGATGAT 976
Db 782 ACAGTGGTCAACGAGCAGTCTTTCGCAATCTGATGATGATGATGATGATGATGATGATGATGAT 841
QY 977 AAAATTTTGTCTTACCGACTTGAAGTAAATGGTCAATAGGCGACGATTTGCGAAGCGA 1036
Db 842 AAAATTTTGTCTTACCGACTTGAAGTAAATGGTCAATAGGCGACGATTTGCGAAGCGA 901
QY 1037 CTCTCTGATCTATTCATGAAGAAATTTGCAACAGCCATTATGAATAGCGACTGTCTAGTCT 1096
Db 902 CTCTCTGATCTATTCATGAAGAAATTTGCAACAGCCATTATGAATAGCGACTGTCTAGTCT 961
QY 1097 TTGACACCGAGCTGACAGCTTTTTCGCAAAATGCGCAATTTTAGGCATCAATGTAACCTA 1156
Db 962 TTGACACCGAGCTGACAGCTTTTTCGCAAAATGCGCAATTTTAGGCATCAATGTAACCTA 1021
QY 1157 TTTTCCATGTGTGAATGGCAATCAAACTTTTCTGCGCAGTGTTTAAACTTTTCACTTTC 1216
Db 1022 TTTTCCATGTGTGAATGGCAATCAAACTTTTCTGCGCAGTGTTTAAACTTTTCACTTTC 1081
```

```
/protein_id="AAC51907.1"
/db_xref="GI:2673966"
/translation="MSRQTATLPTGTSKCPSPQRPVALTGTFTASNNDLASLEPCPVC
SYVLPPIQCQSHLVCSNCRPKLTCCPTCRGLPSIRNLAMKPVANSVLPFCKYAS
FGCEITLPTKEDHDELCEFRPYSCPCPCASCKWQSGSLDAMPHLMHQRKSTTLQ
EDIVFLATINLPGAVDWMMQSCFGFHWLVEKQEKYDGHQFFAIVOLIGTRKQA
ENPFAVELELNGHRRRLTWEATPRS IHGIIATAMNSDCLVFDTSIAQLFAENGLGIN
VTISM"
```



```
QY 1217 ACAGAAAATAGCCACCCATCTGTCTGCCAACCCTAAACCTCTTTTCGGTAGGTGGAAGC 1274
|||||
Db 1082 ACAGAAAATAGCCACCCATCTGTCTGCCAACCCTAAACCTCTTTTCGGTAGGTGGAAGC 1139
|||||

RESULT 16
LOCUS BC042550
DEFINITION Homo sapiens seven in absentia homolog 1 (Drosophila), mRNA (CDNA
clone MGC:34550 IMAGE:5198176), complete cds.
ACCESSION BC042550
VERSION BC042550.1 GI:27503513
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2454)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,I., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Diatelton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richard,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallu,D.E.,
Sutner,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2454)
Strausberg,R.
Direct Submission
Submitted (02-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 50 Row: p Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4506946.

FEATURES
Location/Qualifiers
i. .2454
/organism="Homo sapiens"
/mol_type="mRNA"
```

## gene

```
/db_xref="taxon:9606"
/clone="MGC:34550 IMAGE:5198176"
/tissue_type="Brain, adult, 6 pooled whole brains"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1. .2454
/gene="SIAH1"
/db_xref="LocusID:6477"
/db_xref="MIM:602212"
367..1308
/gene="SIAH1"
/codon_start=1
/product="SIAH1 protein"
/protein_id="AAH42550.1"
/db_xref="GI:27503514"
/db_xref="LocusID:6477"
/db_xref="MIM:602212"
/translation="MTGKATPPSLYSWRGVLTCLPAARTKRKMSQRTATLPTGT
SKCPSPQRPALTGTGTASNNDLASLFEPCVFDYVLPILQCSGHLVCNCRPKLTC
CPTCRGPUGSIRNLAMKXVANSVLPCKKYASSGCEITLPHTEKADHEILCEFRPYSCP
CPGASCKWQGSIDAVMPHLMHQHSITLGGEDIVFLATDNLPGADVMMQSCFGF
HFMVLVEQEKYDGHQFPALVOLLGTRQAEAFYRLNGLNHRRLTWEATPRSIHE
GIATAMNSDCLFDFTSIQAQLFAENGLNGLNITISMIC"
703..1305
/gene="SIAH1"
/note="Sina; Region: Seven in absentia protein family. The
seven in absentia (sina) gene was first identified in
Drosophila. The Drosophila Sina protein is essential for
the determination of the R7 pathway in photoreceptor cell
development: the loss of functional Sina results in the
transformation of the R7 precursor cell to a non- neuronal
cell type. The Sina protein contains an N-terminal RING
finger domain pfam0097. Through this domain, Sina binds
E2 ubiquitin-conjugating enzymes (UbcD1) Sina also
interacts with Tramtrack (TRK88) via PHYL. Tramtrack is a
transcriptional repressor that blocks photoreceptor
determination, while PHYL down-regulates the activity of
TRK88. In turn, the activity of PHYL requires the
activation of the Sevenless receptor tyrosine kinase, a
process essential for R7 determination. It is thought that
thus Sina targets TRK88 for degradation. Therefore
promoting the R7 pathway. Murine and human homologues of
Sina have also been identified. The human homologue Siah-1
also binds E2 enzymes (UbcH5) and through a series of
physical interactions, targets beta-catenin for ubiquitin
degradation. Siah-1 expression is enhanced by p53. itself
promoted by DNA damage. Thus this pathway links DNA damage
to beta-catenin degradation. Sina proteins, therefore,
physically interact with a variety of proteins. The
N-terminal RING finger domain that binds ubiquitin
conjugating enzymes is described in pfam0097, and does
not form part of the alignment for this family. The
remainder C-terminal part is involved in interactions with
other proteins, and is included in this alignment. In
addition to the Drosophila protein and mammalian
homologues, whose similarity was noted previously, this
family also includes putative homologues from
Caenorhabditis elegans, Arabidopsis thaliana"
/db_xref="CDD:pfam03145"
```

## ORIGIN

```
Query Match 75.1% Score 957; DB 9; Length 2454;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 318 AGAATGAGCGCTCAGCTGCTACAGCATTACTACCGGTACTCTGAGTGTCCACATC 377
|||||
Db 456 AGAATGAGCGCTCAGCTGCTACAGCATTACTACCGGTACTCTGAGTGTCCACATC 515
|||||
QY 378 CCAGAGGGTGCTGCTCCCTGACTGGCACAACCTGCATCCACATGACTTGGCCAGTCTTTT 437
|||||
Db 516 CCAGAGGGTGCTGCTCCCTGACTGGCACAACCTGCATCCACATGACTTGGCCAGTCTTTT 575
|||||
```

QY 438 TGAGTGTCCAGTCTGCTTTGACTATGTGTATACGCCCATCTTCTCAATGTACAGTGGCCA 497  
 Db 576 TGAGTGTCCAGTCTGCTTTGACTATGTGTATACGCCCATCTTCTCAATGTACAGTGGCCA 635  
 QY 498 TCTTGTGTTTGTAGCACTGCGCCAAAGCTCAGATGTTCTCAACTTTCGCGGGCCCTTT 557  
 Db 636 TCTTGTGTTTGTAGCACTGCGCCAAAGCTCAGATGTTCTCAACTTTCGCGGGCCCTTT 695  
 QY 558 GGGATCCATTCCGCACTTGGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTCTGTA 617  
 Db 696 GGGATCCATTCCGCACTTGGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTCTGTA 755  
 QY 618 ATATGCGTCTTCTGGATGGAATAAATCTCTGCCACACACAGAAAAAGCAGCATGAAGA 677  
 Db 756 ATATGCGTCTTCTGGATGGAATAAATCTCTGCCACACACAGAAAAAGCAGCATGAAGA 815  
 QY 678 GCTCTGTGAGTTAGGCTTATCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737  
 Db 816 GCTCTGTGAGTTAGGCTTATCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875  
 QY 738 CTCTCTGATGCTGTAATGCCCATCTGATGATCAGCATAGTCCATTACACACCTTACA 797  
 Db 876 CTCTCTGATGCTGTAATGCCCATCTGATGATCAGCATAGTCCATTACACACCTTACA 935  
 QY 798 GGGAGAGGATATAGTTTCTTGTCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCT 857  
 Db 936 GGGAGAGGATATAGTTTCTTGTCTACAGACATTAATCTTCTGCTGCTGCTGCTGCTGCT 995  
 QY 858 GATGATGAGTCTGCTTTTGGCTTCTCACTTCACTGATGATGATGATGATGATGATGATGAT 917  
 Db 996 GATGATGAGTCTGCTTTTGGCTTCTCACTTCACTGATGATGATGATGATGATGATGATGAT 1055  
 QY 918 CGATGTCACACAGCAGTCTTCTGCAATCTGACAGCTGATAGAAACACGCAAGCACTCA 977  
 Db 1056 CGATGTCACACAGCAGTCTTCTGCAATCTGACAGCTGATAGAAACACGCAAGCACTCA 1115  
 QY 978 AAATTTTGTCTACCGACTTGAGCTAAATGGTCTATAGCGACGATTTGACTTGGGAAGCGAC 1037  
 Db 1116 AAATTTTGTCTACCGACTTGAGCTAAATGGTCTATAGCGACGATTTGACTTGGGAAGCGAC 1175  
 QY 1038 TCTCGATCTATTATGAAGGAATTCGACAGGCATTAATGAATAGGACTGTCTAGTCTT 1097  
 Db 1176 TCTCGATCTATTATGAAGGAATTCGACAGGCATTAATGAATAGGACTGTCTAGTCTT 1235  
 QY 1098 TGACACACGATTTGACACGCTTTTTCGAGAAATGSCAATTTAGGCAATCAATGTAACAT 1157  
 Db 1236 TGACACACGATTTGACACGCTTTTTCGAGAAATGSCAATTTAGGCAATCAATGTAACAT 1295  
 QY 1158 TTCCATGTTTGAATGGCAATCAACATTTTTCGCGCAGTGTGTTAAACTTCAGTTTCA 1217  
 Db 1296 TTCCATGTTTGAATGGCAATCAACATTTTTCGCGCAGTGTGTTAAACTTCAGTTTCA 1355  
 QY 1218 CAGAAAAAAGGACCCATCTGTCTGCCAACCTTAAACTCTTTCGCTAGGTGGAAGC 1274  
 Db 1356 CAGAAAAAAGGACCCATCTGTCTGCCAACCTTAAACTCTTTCGCTAGGTGGAAGC 1412

## RESULT 17

HSU63295 1886 bp mRNA linear PRI 09-APR-1998  
 Locus  
 DEFINITION Homo sapiens seven in absentia homolog mRNA, complete cds.  
 ACCESSION U63295  
 VERSION U63295.1 GI:3041824

## KEYWORDS

Source

## ORGANISM

Homo sapiens (human)

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1886)

Nemati, M., Linares-Cruz, G., Bruzzoni-Giovanelli, H., Roperch, J. P.,

Tuynder, M., Bougueleret, L., Cherif, D., Medhioub, M., Pasturaud, P.,

Alvaro, V., Der Sarkissian, H., Cazes, L., Le Paslier, D., Le Gall, I.,

## TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

2 (bases 1 to 1886)

Nemati, M., Linares-Cruz, G., Bruzzoni-Giovanelli, H., Roperch, J. P.,

Tuynder, M., Bougueleret, L., Cherif, D., Medhioub, M., Pasturaud, P.,

Alvaro, V., Der Sarkissian, H., Cazes, L., Le Paslier, D., Le Gall, I.,

Israeli, D., Dausset, J., Sigaux, F., Chumakov, I., Oren, M., Calvo, F.,

Amson, R. B., Cohen, D. and Telerman, A.

Activation of the human homologue of the Drosophila sina gene in

apoptosis and tumor suppression

Proc. Natl. Acad. Sci. U.S.A. 93 (17), 9039-9042 (1996)

96392362

8799150

Submitted (09-JUL-1996) Cancer Research Program, Fondation Jean

Dausset-CEPH, 27, Rue Juliette Dodu, Paris 75010, France

On Apr 9, 1998 this sequence version replaced gi:1508928.

Location/Qualifiers

1..1886

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/chromosome="16"

/map="16q12-q13"

/cell\_type="U937; K562"

/tissue\_type="intestinal epithelium"

1..849

/codon\_start=1

/product="seven in absentia homolog"

/protein\_id="AAC12950.1"

/db\_xref="GI:3041825"

/translation="MSRQTATATPTGTSKPPSRQVPALPTGTASNNDLASLEPCPVC

SGYVLPPLIQSGHLVCSNCRPKLTCCPTCRGPIGSRNLAMEKVAIVSLPCKYAS

FDCEITLPHTEKADHEELCEFRPYSCPCGASCKWQSLDAVPHLMHOKHKTITLQG

EDIVFLADINLPADVVMQSCFHFMLVLEKQYKQDGHQDFAIIVLIGTRQA

ENFATRLSENGHRRRLTWEATPRSIHEGATAIMNSDCLVFDTSIAQLPAENGLIN

VTISM"

Query Match 74.3%; Score 947; DB 9; Length 1886;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 ATGAGCCGTGACAGCTGTACAGCATTACCTACCGGTACCTCGAAGTGTCCACATCCCAG 381

Db 1 ATGAGCCGTGACAGCTGTACAGCATTACCTACCGGTACCTCGAAGTGTCCACATCCCAG 60

QY 382 AGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441

Db 61 AGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120

QY 442 TGTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501

Db 121 TGTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

QY 502 GTTGTAGCACTGTGCGCCCAAGCTCAGTGTGTCACATTTGCGGGGCGCTTTTGGGA 561

Db 181 GTTGTAGCACTGTGCGCCCAAGCTCAGTGTGTCACATTTGCGGGGCGCTTTTGGGA 240

QY 562 TCATTCGCACTTGGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATAT 621

Db 241 TCATTCGCACTTGGCTATGAGAAAGTGGCTAATTCAGTACTTTTCCCTGTAAATAT 300

QY 622 GGTCTCTTGGATGTGAATAACTCTGCCACACACAGAAAAAGCAGACCATCAAGAGCTC 681

Db 301 GGTCTCTTGGATGTGAATAACTCTGCCACACACAGAAAAAGCAGACCATCAAGAGCTC 360

QY 682 TGTGAGTTTAGCCCTATTTCCTGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741

Db 361 TGTGAGTTTAGCCCTATTTCCTGTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 742 CTGGATGCTGTAATCCCCCATCTGATGATCATGCAATAAGTCAATACACCTTACAGGA 801



Db	2411	GTTTAAACTTCAGTTTCACAGAAATAAGGACCCATCTGTCTGCAACCTAAAACTCT	2352
Qy	1259	TTTCGGTAGTGGGAAGC	1274
Db	2351	TTTCGGTAGTGGGAAGC	2336
RESULT 19			
LOCUS	BC018193	2034 bp	mRNA linear PRI 24-SEP-2004
DEFINITION	Homo sapiens seven in absentia homolog 1 (Drosophila), mRNA (cDNA clone IMAGE:3857590), partial cds.		
ACCESSION	BC018193		
VERSION	BC018193.1	GI:17390431	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2034)		
	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickinson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 2034)		
AUTHORS	Director MGC Project.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a> Tissue Procurement: DCTD/DTF cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: <a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a> Contact: <a href="mailto:amg@bcm.tmc.edu">amg@bcm.tmc.edu</a> Gunnarone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAK Plate: 21 Row: b Column: 9 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 40254443.		
FEATURES	Location/Qualifiers 1..2034 /organism="Homo sapiens"		

gene	/mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:3857590" /tissue_type="Ovary, adenocarcinoma" /clone_lib="NIH_MGC_66" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" <1..2034 /gene="SIAH1" /note="synonym: hSIAH1" /db_xref="LocusID:6477" /db_xref="MIM:602212" <1..1061 /gene="SIAH1" /codon_start=3 /product="SIAH1 protein" /protein_id="AAH18193.1" /db_xref="GI:52632427" /db_xref="LocusID:6477" /db_xref="MIM:602212"		
CDS	/translation="APGRGGGAAGGGGRRSALVPGFSEARSPTTMSRQTATAPK TGTSCPPSRVPALTGTASNNDLASLFCVDFYVLPPILOQSGHVLVSNCRPK LTCCTCPGRLGSIIRNLAMEKVANSVLPCKVSSGCEITLPHTEKADHEELCEPRPY SCPCGASCKWQSGSLDAMPHLMOHKSITTLTGEDIVELATDINLPGADVMMQSC PGFHMVLVEKQKDYDQGFPAIVQLIGTRKQAEFAYRLSELGHRRLTWEATPERS IHGIATAIMNSDCLIVFTDTSIAQLFCRKWQFRHCNYFHLVKWQSNIFWPFVKTSVQS KIRHPSVCQPKTSLVSGS"		
ORIGIN	Query Match 65.8%; Score 839; DB 9; Length 2034; Best Local Similarity 99.9%; Pred. No. 0; Matches 958; Conservative 0; Mismatches 0; Indels 1; Gaps 1;		
Qy	317	CAGAATGAGCGCTCAGACTGCTACAGCAATACCTACCGGTACCTCGAAAGTGTCCACCAT	376
Db	100	CAGAATGAGCGCTCAGACTGCTACAGCAATACCTACCGGTACCTCGAAAGTGTCCACCAT	159
Qy	377	CCAGAGGCTGCTCCCTGACTGCGCAACTGCATCCAAATGACATTTGGCGAGTCTTT	436
Db	160	CCAGAGGCTGCTCCCTGACTGCGCAACTGCATCCAAATGACATTTGGCGAGTCTTT	219
Qy	437	TTGAGTGTCCAGTCTGCTTTGACTATGTTACCGCCCACTTCTCAATGTCAGAGTGGCC	496
Db	220	TTGAGTGTCCAGTCTGCTTTGACTATGTTACCGCCCACTTCTCAATGTCAGAGTGGCC	279
Qy	497	ATCTTTGTTGTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCAACTTCGCGGGGCCCTT	556
Db	280	ATCTTTGTTGTAGCAACTGTGCGCCAAAGCTCACATGTTGTCCAACTTCGCGGGGCCCTT	339
Qy	557	TGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTA	616
Db	340	TGGGATCCATTCGCAACTTGGCTATGGAGAAAGTGGCTAAATTCAGTACTTTTCCCTGTA	399
Qy	617	ATATGCGCTTCTCGAGTGTGAATAAATCTTGCACACACAGAAAAGCAGACCATGAAG	676
Db	400	ATATGCGCTTCTCGAGTGTGAATAAATCTTGCACACACAGAAAAGCAGACCATGAAG	459
Qy	677	AGCTCTGTGAGTTTAGGCTTATCTGTCGTCGCTGCTTCTCTGCTAAATGCGCAAG	736
Db	460	AGCTCTGTGAGTTTAGGCTTATCTGTCGTCGCTGCTTCTCTGCTAAATGCGCAAG	519
Qy	737	GCTCTCTGGATGCTGAATGCGCCCATCTGATGCATCAGCATTAAGTCCATTACAACCTTAC	796
Db	520	GCTCTCTGGATGCTGAATGCGCCCATCTGATGCATCAGCATTAAGTCCATTACAACCTTAC	579
Qy	797	AGGAGAGGATATAGTCTTTTCTTGCTACAGCAATTAATCTTCTGCTGCTGTTGACTGGG	856
Db	580	AGGAGAGGATATAGTCTTTTCTTGCTACAGCAATTAATCTTCTGCTGCTGTTGACTGGG	639
Qy	857	TGATCATGCGTCCCTGTTTGGCTTTTCACTTCACTGTTAGTCTTTAGAGAAACAGGAAAAAT	916
Db	640	TGATCATGCGTCCCTGTTTGGCTTTTCACTTCACTGTTAGTCTTTAGAGAAACAGGAAAAAT	699







AUTHORS Spindel,E.R., Pauley,M., Jia,Y., Boyle,N., Jiang,S., Gravett,C., Lupo,S.B., Ali,H., Ojeda,S.R. and Norgren,R.B.  
TITLE Targeted amplification of the 3' end of rhesus macaque orthologs of human genes  
JOURNAL Unpublished (2004)  
COMMENT  
Contact: Spindel ER  
Division of Neuroscience  
Oregon National Primate Research Center  
505 NW 185th Avenue, Beaverton, OR 97006, USA  
Tel: 403-690-5388  
Fax: 503-690-5384  
Email: spindele@ohsu.edu  
Primer A: gaatagcgactgtctagtctt  
Primer B: ttattaccataaaccaatgctc  
STS size: 673  
PCR Profile:  
Hot Start: 95 degrees C for 2.00 min  
Denaturation: 95 degrees C for 0.50 min  
Annealing: 48 degrees C for 0.50 min  
Polymerization: 72 degrees C for 1.00 min  
PCR Cycles: 35  
Extension: 72 degrees C for 7.0 min  
Thermal Cycler: MJ Instruments PTC100  
Protocol:  
Template: 200 ng  
Primer: each 1uM  
dNTP's: each 200 uM  
Tag Polymerase: 0.05 units/ul (Fast Start High Fidelity, Roche)  
Total Vol: 50 ul  
Buffer: MgCl2: 1.8 mM  
Fast Start polymerase reaction buffer (Roche)  
Bases 1-392 are 98% homologous (Blast) to bases 1838-2229 of NM\_003031.2. Bases 405-673 are 95% homologous (Blast) to bases 2231-2489 of NM\_003031.2. Primers were chosen to amplify genomic DNA in the 3' region of SIAH1. As human sequence was used to design the primers, the primer sequences are not included in the rhesus sequence provided below. To obtain additional information regarding primers or clones contact: Dr. Robert Norgren, Dept of Genetics, Cell Biology & Anatomy; University of Nebraska Medical Center; 986395 Nebraska Medical Center; Omaha, NE 68198. Email: rnorgren@unmc.edu  
A database containing sequences associated with this project can be found at: <http://rhesusgenomechip.unomaha.edu/index.html>.  
FEATURES  
source  
1..673  
/organism="Macaca mulatta"  
/mol\_type="genomic DNA"  
/strain="Indian origin"  
/db\_xref="taxon:9544"  
/clone="MMA2345"  
/dev\_stage="Adult"  
/note="Organ: Liver; Vector: pGEM-T Easy; V-type: Plasmid; cDNA amplified from rhesus genomic DNA with the human forward and reverse primers listed above and subcloned into pGEM-T Easy."  
gene  
1..673  
/gene="SIAH1"  
/note="seven in absentia homolog 1"  
STS  
1..673  
/gene="SIAH1"  
ORIGIN  
Query Match 11.1%; Score 142; DB 11; Length 673;  
Best Local Similarity 100.0%; Pred. No. 3.7e-63;  
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1098 TGACACCAGCATGTCACAGCTTTTTCAGAAAAATGGCAATTTAGGCATCAATGTAACATAT 1157

Db 1 TGACACCAGCATGTCACAGCTTTTTCAGAAAAATGGCAATTTAGGCATCAATGTAACATAT 60  
QY 1158 TTCCATGTGTTGAAATGGCAATCAAAACATTTTCTGCCAGTGTGTTTAAACATTTCAAGTTTCA 1217  
Db 61 TTCCATGTGTTGAAATGGCAATCAAAACATTTTCTGCCAGTGTGTTTAAACATTTCAAGTTTCA 120  
QY 1218 CAGAAAAATAAGGCACCCATCTG 1239  
Db 121 CAGAAAAATAAGGCACCCATCTG 142  
RESULT 25  
LOCUS CQ511941 417 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 43808 from Patent WO0160860.  
ACCESSION CQ511941  
VERSION CQ511941.1 GI:41478205  
KEYWORDS Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 Schlegel,R., Endege,W.O. and Monahan,J.E.  
AUTHORS Genes differentially expressed in human prostate cancer and their  
TITLE use  
JOURNAL Patent: WO 0160860-A 43808 23-AUG-2001;  
MILLENNIUM PREDICTIVE MEDICINE, INC. (US)  
FEATURES  
source  
1..417  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 10.9%; Score 139; DB 6; Length 417;  
Best Local Similarity 100.0%; Pred. No. 1.5e-61;  
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1136 ATTATAGGCATCAATGTAACATTTTCCATGTGTTGAAATGGCAATCAAAACATTTCTGGCC 1195  
Db 413 ATTATAGGCATCAATGTAACATTTTCCATGTGTTGAAATGGCAATCAAAACATTTCTGGCC 354  
QY 1136 AGTGTGTTAAAACTTCAGTTTCACAGAAAAATAGGCACCCATCTGTCGCCAACCTTAAAC 1255  
Db 353 AGTGTGTTAAAACTTCAGTTTCACAGAAAAATAGGCACCCATCTGTCGCCAACCTTAAAC 294  
QY 1256 TCTTTCGGTAGTGGAGC 1274  
Db 293 TCTTTCGGTAGTGGAGC 275  
RESULT 26  
LOCUS CQ516613 473 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 48480 from Patent WO0160860.  
ACCESSION CQ516613  
VERSION CQ516613.1 GI:41482877  
KEYWORDS Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 Schlegel,R., Endege,W.O. and Monahan,J.E.  
AUTHORS Genes differentially expressed in human prostate cancer and their  
TITLE use  
JOURNAL Patent: WO 0160860-A 48480 23-AUG-2001;  
MILLENNIUM PREDICTIVE MEDICINE, INC. (US)  
FEATURES  
source  
1..473  
/organism="Homo sapiens"



```
JOURNAL Patent: WO 0160860-A 34958 23-AUG-2001;
FEATURES Millennium Predictive Medicine, Inc. (US)
SOURCE Location/Qualifiers
1. .441
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 10.0%; Score 128; DB 6; Length 473;
Best Local Similarity 100.0%; Pred. No. 9.8e-56;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 969 GCAAGCTGAAATTTTGTCTTACCGACTTGAGCTTAATGGTCATAGCGAGATGACTTG 1028
|
|
|
Db 473 GCAAGCTGAAATTTTGTCTTACCGACTTGAGCTTAATGGTCATAGCGAGATGACTTG 414
|
|
|
QY 1029 GGAAGCGACTCTCGATCTATTTCATGAGCAATGCAAGCCATTATGATGAGCACTG 1088
|
|
|
Db 413 GGAAGCGACTCTCGATCTATTTCATGAGCAATGCAAGCCATTATGATGAGCACTG 354
|
|
|
QY 1089 TCTAGTCT 1096
|
|
|
Db 353 TCTAGTCT 346
|
|
|

RESULT 27
CO481952/c 411 bp DNA linear PAT 30-JAN-2004
LOCUS
DEFINITION Sequence 13819 from Patent WO0160860.
ACCESSION CO481952
VERSION CO481952.1 GI:41447571
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 13819 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES Location/Qualifiers
SOURCE 1. .411
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 9.3%; Score 118; DB 6; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-50;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 TTTCATGTGTTGAATGCAATCAAACTTTCTGCGCCAGTGTTTAAACTTCAGTTTC 1216
|
|
|
Db 364 TTTCATGTGTTGAATGCAATCAAACTTTCTGCGCCAGTGTTTAAACTTCAGTTTC 305
|
|
|
QY 1217 ACAGAAATAAGGACCCATCTCTGCGCAACTTAACTTTTCGGTAGGTGGAAGC 1274
|
|
|
Db 304 ACAGAAATAAGGACCCATCTCTGCGCAACTTAACTTTTCGGTAGGTGGAAGC 247
|
|
|

RESULT 28
CQ503091/c 441 bp DNA linear PAT 30-JAN-2004
LOCUS
DEFINITION Sequence 34958 from Patent WO0160860.
ACCESSION CQ503091
VERSION CQ503091.1 GI:41468727
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
```

```
JOURNAL Patent: WO 0160860-A 34958 23-AUG-2001;
FEATURES Millennium Predictive Medicine, Inc. (US)
SOURCE Location/Qualifiers
1. .441
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 9.3%; Score 118; DB 6; Length 441;
Best Local Similarity 100.0%; Pred. No. 1.9e-50;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 TTTCATGTGTTGAATGCAATCAAACTTTCTGCGCCAGTGTTTAAACTTCAGTTTC 1216
|
|
|
Db 399 TTTCATGTGTTGAATGCAATCAAACTTTCTGCGCCAGTGTTTAAACTTCAGTTTC 340
|
|
|
QY 1217 ACAGAAATAAGGACCCATCTCTGCGCAACTTAACTTTTCGGTAGGTGGAAGC 1274
|
|
|
Db 339 ACAGAAATAAGGACCCATCTCTGCGCAACTTAACTTTTCGGTAGGTGGAAGC 282
|
|
|

RESULT 29
BV102520 523 bp DNA linear STS 28-FEB-2004
LOCUS
DEFINITION MARC 14423-14424:1027951438:1 SCF - porcine spleen Sus scrofa STS
genomic, sequence tagged site.
ACCESSION BV102520
VERSION BV102520.1 GI:44680182
KEYWORDS STS.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 523)
Freking,B.A., Nonneman,D.P., Rohrer,G.A., Fahrenkrug,S.C.,
Smith,T.P.L. and Keele,J.W.
TITLE Development of a comprehensive SNP-based comparative linkage map
for pigs
JOURNAL Unpublished (2004)
COMMENT
Contact: Freking BA
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4278
Fax: 402 762 4173
Email: freking@email.marc.usda.gov
Primer A: ATGCTCTCTGGATGAGAGA
Primer B: CGCGCTATGACCATTTAG
STS size: 400
PCR Profile:
Hotstart: 95 degrees for 15 minutes
Denature: 95 degrees for 30 seconds
Anneal: 56 degrees
Extension: 68 degrees for 2 minutes
Cycles: 32 to 45
Protocol:
Template: 50-200 ng genomic DNA
Primer: each 20 pmoles
dNTPs: each 88 uM
Tag Polymerase: 0.25 units (Qiagen HotStar)
Buffer: Commercially supplied Qiagen HotStar buffer

The STS is derived from PCR amplicons generated from genomic DNA,
sequenced from each end using the amplification primers. The
sequence does not necessarily represent the entire amplicon.
Sequence derived from PolyPhred was trimmed from each end of each
unique contig until five consecutive bases exceeded a quality score
threshold of 20, and the next ten bases averaged a quality score of
20 or greater. Amplicon size was estimated by agarose gel
electrophoresis.
```

```
FEATURES
  source
    Location/Qualifiers
      1..523
        /organism="Sus scrofa"
        /mol_type="genomic DNA"
        /strain="white composite, duroc, meishan, minzhu,
        fengjing, crossbreds"
        /db_xref="taxon:9823"
        /sex="male and female"
        /clone_lib="SCF - porcine spleen"
        /dev_stage="adult"
        /note="Organ: spleen"
      124..523
        primer_bind
          primer_bind
            complement(504..523)
      ORIGIN
        Query Match
          7.2%; Score 92; DB 11; Length 523;
          Best Local Similarity 98.8%; Pred. No. 1.1e-36;
          Matches 242; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
          QY 652 CACACAGAAAAGCAGACCATCAAGAGCTCTGTGAGTTTAGGCTTATTCCTGTCCGTGC 711
          Db 156 CACACAGAAAAGCAGACCATCAAGAGCTCTGTGAGTTTAGGCTTATTCCTGTCCGTGC 215
          QY 712 CTGTGCTCTCTGTAATGCGAAGCTCTCTGGATGCTGTAAATGCCCATCTGATGCAT 771
          Db 216 CTGTGCTCTCTGTAATGCGAAGCTCTCTGGATGCTGTAAATGCCCATCTGATGCAT 275
          QY 772 GAGCATAAGTTCATTACACACCTACAGGAGGAGATATAGTTTTCTTCTACAGACATT 831
          Db 276 GAGCATAAGTTCATTACACACCTACAGGAGGAGATATAGTTTTCTTCTACAGACATT 335
          QY 832 AATCTTCTGTGCTGTTCTACCTGGGTGATGATGCTGTTTGGCTTTTCACCTTCATG 891
          Db 336 AATCTTCTGTGCTGTTCTACCTGGGTGATGATGCTGTTTGGCTTTTCACCTTCATG 395
          QY 892 TTAGT 896
          Db 396 TTAGT 400
      RESULT 30
      AC133654
      LOCUS
      DEFINITION
        Mus musculus BAC clone RP24-332N10 from chromosome 15, complete
        sequence.
      ACCESSION
        AC133654
      VERSION
        AC133654.4 GI:37361082
      KEYWORDS
        HTG.
      SOURCE
        Mus musculus
        (house mouse)
      ORGANISM
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
      REFERENCE
        1 (bases 1 to 194622)
        Tomlinson,C.
      AUTHORS
        The sequence of Mus musculus BAC clone RP24-332N10
      JOURNAL
        Unpublished (2001)
      REFERENCE
        2 (bases 1 to 194622)
        Wilson,R.
      AUTHORS
        Sequencing of Mus musculus
      JOURNAL
        Unpublished (2001)
      REFERENCE
        3 (bases 1 to 194622)
        McPherson,J.D. and Waterston,R.H.
      AUTHORS
        Direct Submission
      JOURNAL
        Submitted (16-SEP-2002) Genome Sequencing Center, 4444 Forest Park
        Parkway, St. Louis, MO 63108, USA
      REFERENCE
        4 (bases 1 to 194622)
        Wilson,R.K.
      AUTHORS
        Direct Submission
      JOURNAL
        Submitted (14-AUG-2003) Genome Sequencing Center, 4444 Forest Park
        Parkway, St. Louis, MO 63108, USA
      REFERENCE
        5 (bases 1 to 194622)
        Wilson,R.K.
      AUTHORS
        Direct Submission
      JOURNAL
        Submitted (02-OCT-2003) Genome Sequencing Center, 4444 Forest Park
        Parkway, St. Louis, MO 63108, USA
      REFERENCE
        6 (bases 1 to 194622)
        Wilson,R.
      AUTHORS
        Direct Submission
      JOURNAL
        Submitted (27-NOV-2003) Department of Genetics, Washington
        University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
      COMMENT
        On Oct 2, 2003 this sequence version replaced gi:33636837.
        -----
        Center: Washington University Genome Sequencing Center
        Center code: WUGSC
        Web site: http://genome.wustl.edu
        Contact: submissions@watson.wustl.edu
        -----
        Summary Statistics
        -----
        Center project name: M_BB0332N10
        -----
        NOTICE: This sequence may not represent the entire insert of this
        clone. It may be shorter because we only sequence overlapping
        clone sections once, or longer because we provide a small overlap
        between neighboring data submissions.
        This sequence was finished as follows unless otherwise noted:
        all regions were double stranded, sequenced with an alternate
        chemistry, or covered by high quality data (i.e.; phred quality >=
        30); an attempt was made to resolve all sequencing problems, such
        as compressions and repeats; all regions were covered by sequence
        from more than one subclone; and the assembly was confirmed by
        restriction digest.
        MAPPING INFORMATION:
        Mapping information for this clone was provided by Dr. Wes Warren,
        Department of Genetics, Washington University, St. Louis MO. For
        additional information about the map position of this sequence, see
        http://genome.wustl.edu
        SOURCE INFORMATION:
        The RPCI-24 BAC Library has been constructed by Pieter de Jong and
        coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
        and/or brain genomic DNA. The clone and detailed information can be
        obtained from Pieter de Jong and coworkers at http://www.chori.org
        NEIGHBORING SEQUENCE INFORMATION:
        This sequence is the entire insert of the clone.
        Location/Qualifiers
          1..194622
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /chromosome="15"
            /map="15"
            /clone="RP24-332N10"
            /clone_lib="RPCI-24"
            /rpt_family="MaLR"
            143..1005
            /rpt_family="B4"
            1020..1158
            /rpt_family="B4"
            2576..2612
            /rpt_family="ERVK"
            2613..2778
            /rpt_family="RMR19B"
            2779..3057
            /rpt_family="ERVK"
            3055..3297
            /rpt_family="MaLR"
            4466..4807
            /rpt_family="MaLR"
            5016..5225
            /rpt_family="B4"
            5382..5451
            /rpt_family="B4"
            5603..5872
```

```

/rpt family="L1"
repeat_region 6054..6149
/rpt family="Alu"
repeat_region 6240..6368
/rpt family="B2"
repeat_region 6377..6503
/rpt family="ERV1"
repeat_region 10251..10300
/rpt family="ERV1"
repeat_region 10350..10463
/rpt family="Alu"
repeat_region 10908..11077
/rpt family="B2"
repeat_region 11475..11634
/rpt family="RMR10A"
repeat_region 11826..11944
/rpt family="Alu"
repeat_region 12158..12280
/rpt family="ID"
repeat_region 12281..12467
/rpt family="B2"
repeat_region 12468..12498
/rpt family="ID"
repeat_region 13303..13391
/rpt family="ID"
complement(13318..13388)
/product="tRNA-Val"
/note="likely pseudogene (HMM Sc=27.16 / Sec struct
Sc=-6.26)"

repeat_region 13780..13967
/rpt family="B2"
repeat_region 14286..14400
/rpt family="Alu"
repeat_region 14526..14927
/rpt family="MaLR"
repeat_region 15873..16228
/rpt family="MaLR"
repeat_region 16241..16459
/rpt family="B2"
repeat_region 17324..17611
/rpt family="MaLR"
repeat_region 18038..18186
/rpt family="Alu"
repeat_region 18484..18588
/rpt family="Alu"
repeat_region 19067..19249
/rpt family="B2"
repeat_region 19263..19473
/rpt family="L1"
repeat_region 19531..19718
/rpt family="B2"
repeat_region 20889..21104
/rpt family="MaLR"
repeat_region 21656..21794
/rpt family="Alu"
repeat_region 22024..22178
/rpt family="B2"
repeat_region 22205..22386
/rpt family="B2"
repeat_region 22470..22612
/rpt family="Alu"
repeat_region 22975..23153
/rpt family="B2"
repeat_region 23561..23635
/rpt family="ID"
repeat_region 23927..24076
/rpt family="Alu"
repeat_region 24624..24748
/rpt family="Alu"
repeat_region 27545..27621
/rpt family="B4"
repeat_region 27622..27716
/rpt family="Alu"

Query Match 6.9%; Score 88; DB 10; Length 194622;
Best Local Similarity 100.0%; Pred. No. 7.5e-35;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1096 TTTCACACGAGTCACAGCTTTTTCAGAAAATGGCAATTTAGGCATCAATGTAAC 1155
Db 142060 TTTCACACGAGTCACAGCTTTTTCAGAAAATGGCAATTTAGGCATCAATGTAAC 142119

QY 1156 ATTTCATGTGTGAAATGGCAATCAAA 1183
Db 142120 ATTTCATGTGTGAAATGGCAATCAAA 142147

RESULT 31
CQ472783/c
LOCUS CQ472783 438 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 4650 from Patent WO0160860.
ACCESSION CQ472783
VERSION CQ472783.1 GI:41438402
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNAL Patent: WO 0160860-A 4650 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source
Location/Qualifiers
1..438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 6.7%; Score 85; DB 6; Length 438;
Best Local Similarity 100.0%; Pred. No. 5.4e-33;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 CTGGCCAGTGTAAACCTTCAGTTTCAGAAAATAGGCACCCATCTGTCTGCCAAC 1249
Db 366 CTGGCCAGTGTAAACCTTCAGTTTCAGAAAATAGGCACCCATCTGTCTGCCAAC 307

QY 1250 TAAACTCTTCGGTAGGTGGAGC 1274
Db 306 TAAACTCTTCGGTAGGTGGAGC 282

```

```
RESULT 32
BC046317
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
12477932
2 (bases 1 to 1645)
Direct Submission
Submitted (31-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Contact: nisc.mgc@hgrl.nih.gov
Web site: http://www.nisc.nih.gov/
Ahter.N., Avele.K., Beckstrom-Sternberg,S.M., Benjamin.B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dierich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 100 Row: e Column: 20.
Location/Qualifiers
1..1645
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:54763 IMAGE:6314641"
FEATURES
source
1..1645
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:54763 IMAGE:6314641"
BC046317 1645 bp mRNA linear ROD 30-JUN-2004
Mus musculus seven in absentia 1A, mRNA (cdna clone MGC:54763
IMAGE:6314641), complete cds.
BC046317
MGC.
BC046317.1 GI:28277393
Mus musculus (house mouse)
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1645)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uscin,T.B., Toshiyuki,S.,
Carrinci,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smalhus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1645)
Strausberg,R.
Direct Submission
Submitted (31-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Contact: nisc.mgc@hgrl.nih.gov
Web site: http://www.nisc.nih.gov/
Ahter.N., Avele.K., Beckstrom-Sternberg,S.M., Benjamin.B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dierich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 100 Row: e Column: 20.
Location/Qualifiers
1..1645
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:54763 IMAGE:6314641"
/tissue_type="Olfactory epithelium, neonatal mouse,
C57Bl/6"
/clone_lib="NIH MGC_129"
/lab_host="DHIOB"
/note="Vector: PCMV-SPORT6"
1..1645
/gene="Siahlia"
/db_xref="LocusID:20437"
/db_xref="MGI:108064"
193..1041
/gene="Siahlia"
/codon_start=1
/product="seven in absentia 1A"
/protein_id="AAH46317.1"
/db_xref="GI:28277394"
/db_xref="LocusID:20437"
/db_xref="MGI:108064"
/translation="MSRQTATALPTGTSKCPSPQRPALPTGTASNNDLASLFCPCV
FDVLPPILOQSGHLVCSNCRPKLTCCPTGCRPLGIRNLAMEKVASVLPFCCKYAS
SGCEITLPHTEKAHEELCEPRPSCPCGASCKWQSGSLDAVMPHLMHQHKSITTLQ
EDIVFLATDINLPAGADVMMQSCFGPHFMLVLEKQBYDGHQOPFAIVOLIGTKOA
ENPAYRELNGHRRRLTWEATPRSHIEGIAINWNSDCLVPTDTSIAQLFAENGNLGIN
VTISM"
ORIGIN
Query Match 6.0%; Score 77; DB 10; Length 1645;
Best Local Similarity 100.0%; Pred. No. 8.1e-29;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1096 TTTGACACGAGCATTCGACAGCTTTTTCAGAAAATGCAATTTAGGCATCAATGTAAC 1155
Db 967 TTTGACACGAGCATTCGACAGCTTTTTCAGAAAATGCAATTTAGGCATCAATGTAAC 1026
QY 1156 ATTTCCATGTGTGAAA 1172
Db 1027 ATTTCCATGTGTGAAA 1043
RESULT 33
AX305603
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1968
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 6.0%; Score 77; DB 6; Length 1968;
Best Local Similarity 100.0%; Pred. No. 7.9e-29;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1096 TTTGACACGAGCATTCGACAGCTTTTTCAGAAAATGCAATTTAGGCATCAATGTAAC 1155
Db 895 TTTGACACGAGCATTCGACAGCTTTTTCAGAAAATGCAATTTAGGCATCAATGTAAC 954
QY 1156 ATTTCCATGTGTGAAA 1172
Db 955 ATTTCCATGTGTGAAA 971
```

```

RESULT 34
MMSIAHIA
LOCUS      MMSIAHIA      1968 bp      mRNA      linear      ROD 03-NOV-1993
DEFINITION M.musculus siah-1A protein mRNA.
ACCESSION  Z19579
VERSION     Z19579.1  GI:297034
KEYWORDS    siah-1A protein.
SOURCE      Mus musculus (house mouse)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 1968)
AUTHORS     Della,N.G., Senior,P.V. and Bowtell,D.D.
TITLE       Isolation and characterisation of murine homologues of the
            Drosophila seven in absentia gene (sina)
JOURNAL     Development 117 (4), 1333-1343 (1993)
MEDLINE     94008536
PUBMED      8404535
REFERENCE   2 (bases 1 to 1968)
AUTHORS     Bowtell,D.D.
TITLE       Direct Submission
JOURNAL     Submitted (15-JAN-1993) David D.L. Bowtell, Developmental Biology
            Group, Howard Florey, Institute of Exp. Physiology and Medicine,
            University of, Melbourne, Parkville, Victoria, 3052, Australia
            Location/Qualifiers
FEATURES    source
            1..1968
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="random bred Swiss"
               /db_xref="taxon:10090"
            1..120
            121..969
               /codon_start=1
               /product="siaH-1A protein"
               /protein_id="CAA79630.1"
               /db_xref="GI:297035"
               /db_xref="GOA:Q06984"
               /translation="MSROTATALPTGTCPCPSQRPALTTGTASNNDLASLFCPCV
               FDYVLPILQCSQHLVCSNCRPKLTCCPTCPGLGSIRNLAMEKVANSVLPCKVAS
               SCCEITLPHTEAHEELCEFPYSCPCPGACKQGSIDAVPHLMHQHKSITTLQ
               EDIVFLATDINLPGAVWQSCFGFHWLEKQEKDGHQGFALVQLLGTTRQA
               ENFATLDELNGHRRRLTWEATPRSIHEGIATMNSDCLVFDTSIAQLFAENGILGIN
               VTISM"
            970..1968
            1491..1496
            1515
            1943..1948
            1968
ORIGIN
3'UTR
polyA_signal
polyA_site
polyA_signal
polyA_site
1968

Query Match      6.0%; Score 77; DB 10; Length 1968;
Best Local Similarity 100.0%; Pred. No. 7.9e-29;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1096 TTTGACACGAGATTGCACAGCTTTTTCAGAGAAATGGCAATTAGGCATCAATGTAAC 1155
      |||||||
Db 895 TTTGACACGAGATTGCACAGCTTTTTCAGAGAAATGGCAATTAGGCATCAATGTAAC 954
      |||||||

QY 1156 ATTTCATGTGTTGAAA 1172
      |||||||
Db 955 ATTTCATGTGTTGAAA 971
      |||||||

RESULT 35
AC142211
LOCUS      Mus musculus BAC clone RP24-501B23 from chromosome 8, complete
DEFINITION sequence.
ACCESSION  AC142211
VERSION     AC142211.3  GI:48675474
KEYWORDS    HTG.

```

```

SOURCE
ORGANISM    Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 175636)
AUTHORS     Cotton,M., Haakenson,W. and Shahid,S.
TITLE       The sequence of Mus musculus BAC clone RP24-501B23
JOURNAL     Unpublished (2001)
REFERENCE   2 (bases 1 to 175636)
AUTHORS     McPherson,J.D. and Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (25-MAR-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3 (bases 1 to 175636)
AUTHORS     McPherson,J.D. and Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (26-MAR-2003) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   4 (bases 1 to 175636)
AUTHORS     Wilson,R.K.
TITLE       Direct Submission
JOURNAL     Submitted (12-JUN-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   5 (bases 1 to 175636)
AUTHORS     Wilson,R.K.
TITLE       Direct Submission
JOURNAL     Submitted (10-JUL-2004) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jun 12, 2004 this sequence version replaced gi:29244764.
            ----- Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            ----- Summary Statistics
            -----
            Center project name: M_BB0501B23
            -----

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RPCI-24 BAC Library has been constructed by Pieter de Jong and
coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.
Location/Qualifiers
1..175636
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/maps="8"
/clone="RP24-501B23"
/clone_lib="RPCI-24"
repeat_region 5..754

```

```
repeat_region /rpt_family="L1"
898. .1045
repeat_region /rpt_family="B4"
1156. .1421
trna /rpt_family="L1"
2845. .2917
/product="tRNA-Ser"
/notes="Likely pseudogene (HMM Sc=34.44 / Sec struct
Sc=-9.72)"
repeat_region 2848. .3027
/rpt_family="B2"
repeat_region 3043. .3183
/rpt_family="L1"
repeat_region 3901. .4186
/rpt_family="MaLR"
repeat_region 4637. .4953
/rpt_family="MaLR"
repeat_region 4958. .5210
/rpt_family="MaLR"
repeat_region 5211. .5580
/rpt_family="MaLR"
repeat_region 5563. .5684
/rpt_family="MaLR"
repeat_region 5685. .5776
/rpt_family="MaLR"
repeat_region 5708. .5780
/rpt_family="Alu"
repeat_region 6154. .6358
/rpt_family="B2"
repeat_region 6489. .6669
/rpt_family="B2"
repeat_region 6671. .6864
/rpt_family="B2"
repeat_region 6865. .6878
/rpt_family="B4"
repeat_region 6879. .7027
/rpt_family="Alu"
repeat_region 7028. .7088
/rpt_family="B4"
repeat_region 7680. .7883
/rpt_family="B2"
repeat_region 8256. .8757
/rpt_family="ERVL"
repeat_region 9203. .9293
/rpt_family="MIR"
repeat_region 10285. .10327
/rpt_family="MIR"
repeat_region 10331. .10827
/rpt_family="ERVK"
repeat_region 11819. .12024
/rpt_family="MaLR"
repeat_region 12030. .12203
/rpt_family="B2"
repeat_region 12320. .12669
/rpt_family="MaLR"
repeat_region 12730. .12839
/rpt_family="Alu"
repeat_region 12897. .13082
/rpt_family="MaLR"
repeat_region 13526. .13593
/rpt_family="MaLR"
repeat_region 13750. .13925
/rpt_family="MaLR"
repeat_region 13941. .14158
/rpt_family="B2"
repeat_region 14436. .14584
/rpt_family="L1"
repeat_region 15670. .15743
/rpt_family="ID"
repeat_region 16031. .16083
/rpt_family="MIR"
repeat_region 16332. .16525
/rpt_family="B4"

repeat_region 16578. .16719
/rpt_family="Alu"
repeat_region 16829. .17245
/rpt_family="L2"
repeat_region 17253. .17350
/rpt_family="B4"
repeat_region 17578. .17727
/rpt_family="L2"
repeat_region 17982. .18175
/rpt_family="B2"
repeat_region 18489. .18769
/rpt_family="ERVK"
repeat_region 18987. .19304
/rpt_family="MaLR"
repeat_region 19694. .19900
/rpt_family="B2"
repeat_region 20140. .20268
/rpt_family="Alu"
repeat_region 20369. .20579
/rpt_family="MaLR"
repeat_region 20599. .20725
/rpt_family="Alu"
repeat_region 21080. .21149
/rpt_family="ID"
repeat_region 21348. .21457
/rpt_family="Alu"
repeat_region 22354. .22573
/rpt_family="B2"
repeat_region 22986. .23157
/rpt_family="MaLR"
repeat_region 23168. .23278
/rpt_family="B4"
repeat_region 23456. .23520
/rpt_family="MaLR"
repeat_region 23996. .24325
/rpt_family="MaLR"
repeat_region 25121. .25290
/rpt_family="B4"
repeat_region 25402. .25492
/rpt_family="Alu"
repeat_region 25698. .25886
/rpt_family="L1"
repeat_region 26091. .26240
/rpt_family="L1"
repeat_region 26241. .26382
/rpt_family="Alu"
repeat_region 26383. .26516
/rpt_family="L1"

Query Match 6.0%; Score 77; DB 10; Length 175636;
Best Local Similarity 100.0%; Pred.No. 5e-29;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1096 TTGCACACGAGTCACAGCTTTTCAGAAAATGCAATTTAGGCATCAATGTAAC 1155
|||||
Db 97733 TTGCACACGAGTCACAGCTTTTCAGAAAATGCAATTTAGGCATCAATGTAAC 97792
|||||

QY 1156 ATTTCATGTGTGAAA 1172
|||||
Db 97793 ATTTCATGTGTGAAA 97809
|||||

RESULT 36
LOCUS BV104504
DEFINITION MARC 14423-14424:1010076472:1 RTS-1 Bos indicus x Bos taurus STS
ACCESSION BV104504
VERSION BV104504.1 GI:45182422
KEYWORDS STS.
SOURCE Bos indicus x Bos taurus (hybrid cattle)
ORGANISM Bos indicus x Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

ORGANISM Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovinae; Bos.  
REFERENCE 1 (bases 1 to 732)  
AUTHORS Snelling,W.M., Casas,E., Stone,R.T., Smith,T.P.L., Keele,J.W. and  
Bennett,G.L.  
TITLE Linkage mapping bovine ESTs  
JOURNAL Unpublished (2004)  
COMMENT

Contact: Stone RT  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4166  
Fax: 402 762 4173  
Email: stone@mail.marc.usda.gov  
Primer A: ATGGCTCTCTGGATGTGAGA  
Primer B: CGCCGCTATGACCATTTAG  
STS size: 400  
PCR Profile:

Hotstart: 95 degrees for 15 minutes  
Denature: 95 degrees for 30 seconds  
Anneal: 58 degrees for 30 seconds  
Extension: 68 degrees for 1 minute  
Cycles: 32 to 45

Protocol:  
Template: 25-50 ng genomic DNA  
Primer: each 0.4 uM  
dNTPs: each 100 uM  
Taq Polymerase: 0.25 units (Qiagen HotStar)

Buffer: Commercially supplied Qiagen HotStar buffer

The STS is derived from PCR amplicons generated from genomic DNA, sequenced from each end using the amplification primers. The sequence does not necessarily represent the entire amplicon. Sequence derived from PolyPhred was trimmed from each end of each unique contig until five consecutive bases exceeded a quality score threshold of 20, and the next ten bases averaged a quality score of 20 or greater. Amplicon size was estimated by agarose gel electrophoresis.

#### FEATURES

Location/Qualifiers  
1..732  
/organism="Bos indicus x Bos taurus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:30522"  
/clone\_lib="RTS-1"  
/dev\_stage="adult"  
/note="taurus and indicus crossbreds"  
<1..>732

#### STS

##### ORIGIN

Query Match 4.9%; Score 62; DB 11; Length 732;  
Best Local Similarity 100.0%; Pred. No. 7.5e-21;  
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 808 ATAGTTTCTTCTGCTACAGACATTAACTCTCTGGTGTGCTGGTGATGATGCAG 867  
Db 521 ATAGTTTCTTCTGCTACAGACATTAACTCTCTGGTGTGCTGGTGATGATGCAG 580  
QY 868 TC 869  
Db 581 TC 582

#### RESULT 37

BV104504/c 732 bp DNA linear STS 05-MAR-2004  
LOCUS MARC 14423-14424:1010076472:1 RTS-1 Bos indicus x Bos taurus STS  
DEFINITION genomic, sequence tagged site.  
ACCESSION BV104504  
VERSION BV104504.1 GI:45182422  
KEYWORDS STS.  
SOURCE Bos indicus x Bos taurus (hybrid cattle)

ORGANISM Bos indicus x Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eumalia; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovinae; Bos.  
REFERENCE 1 (bases 1 to 732)  
AUTHORS Snelling,W.M., Casas,E., Stone,R.T., Smith,T.P.L., Keele,J.W. and  
Bennett,G.L.  
TITLE Linkage mapping bovine ESTs  
JOURNAL Unpublished (2004)  
COMMENT

Contact: Stone RT  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4166  
Fax: 402 762 4173  
Email: stone@mail.marc.usda.gov  
Primer A: ATGGCTCTCTGGATGTGAGA  
Primer B: CGCCGCTATGACCATTTAG  
STS size: 400  
PCR Profile:

Hotstart: 95 degrees for 15 minutes  
Denature: 95 degrees for 30 seconds  
Anneal: 58 degrees for 30 seconds  
Extension: 68 degrees for 1 minute  
Cycles: 32 to 45

Protocol:  
Template: 25-50 ng genomic DNA  
Primer: each 0.4 uM  
dNTPs: each 100 uM  
Taq Polymerase: 0.25 units (Qiagen HotStar)

Buffer:

Commercially supplied Qiagen HotStar buffer

The STS is derived from PCR amplicons generated from genomic DNA, sequenced from each end using the amplification primers. The sequence does not necessarily represent the entire amplicon. Sequence derived from PolyPhred was trimmed from each end of each unique contig until five consecutive bases exceeded a quality score threshold of 20, and the next ten bases averaged a quality score of 20 or greater. Amplicon size was estimated by agarose gel electrophoresis.

#### FEATURES

Location/Qualifiers  
1..732  
/organism="Bos indicus x Bos taurus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:30522"  
/clone\_lib="RTS-1"  
/dev\_stage="adult"  
/note="taurus and indicus crossbreds"  
<1..>732

#### STS

##### ORIGIN

Query Match 4.9%; Score 62; DB 11; Length 732;  
Best Local Similarity 100.0%; Pred. No. 7.5e-21;  
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 808 ATAGTTTCTTCTGCTACAGACATTAACTCTCTGGTGTGCTGGTGATGATGCAG 867  
Db 192 ATAGTTTCTTCTGCTACAGACATTAACTCTCTGGTGTGCTGGTGATGATGCAG 133  
QY 868 TC 869  
Db 132 TC 131

#### RESULT 38

AB067814 1457 bp mRNA linear ROD 31-OCT-2001  
LOCUS Rattus norvegicus Shiahla mRNA for shiahla protein, complete cds.  
DEFINITION AB067814  
ACCESSION AB067814  
VERSION AB067814.1 GI:16549114  
KEYWORDS



SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS Yamaguchi, A., Hori, O. and Tohyama, M.  
TITLE rat Slahla  
JOURNAL Unpublished  
2 (bases 1 to 1457)  
AUTHORS Yamaguchi, A., Hori, O. and Tohyama, M.  
TITLE Direct Submission  
JOURNAL Submitted (05-AUG-2001) Ateushi Yamaguchi, Graduate School of  
Medicine, Osaka University, Anatomy and Neuroscience; 2-2,  
Yamada-oka, Suita, Osaka 565-0871, Japan  
(E-mail: yama@anat2.med.osaka-u.ac.jp, Tel: 81-6-6879-3221 (ex. 3221),  
Fax: 81-6-6879-3229)

FEATURES  
source Location/Qualifiers  
1. .1457  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
1. .1457  
/gene="Slahla"  
19. .888  
/gene="Slahla"  
/codon\_start=1  
/product="slahla protein"  
/protein\_id="BAB70753.1"  
/db\_xref="GI:16549115"  
/translation="MCRLEMSRQTALPTGTSKPPSQRPALPTGTSANNNDLAS  
LPCKVASSCEITLPHTEKAEHEELCEFRPYSCPGASCKWQSGSLDAMPHMVRQHK  
SITTLQGEDIVFLATDINLPGADVVRMQSCFGEFHMVLVQKEXYGGHQFFAFVQL  
IGTRKQENFAIVELNGHRRRLTWEATPRSIHEGIATAIMNSDCLVFTSIAQLFAE  
NGLNGINVTISM"

ORIGIN  
Query Match 4.2%; Score 53; DB 10; Length 1457;  
Best Local Similarity 100.0%; Pred. No. 4e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1120 TTTGCAGAAATGGCAATTAGGCATCAATGTAACATTTCCATGTGTTGAAA 1172  
|||||  
Db 838 TTTGCAGAAATGGCAATTAGGCATCAATGTAACATTTCCATGTGTTGAAA 890  
|||||

RESULT 39  
AF389476 1465 bp mRNA linear ROD 20-MAR-2002  
LOCUS Rattus norvegicus SIAH-1A mRNA, complete cds.  
DEFINITION AF389476  
ACCESSION AF389476  
VERSION AF389476.1 GI:19550382  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS Wheeler, T.C., Chin, L.S., Li, Y., Roudabush, F.L. and Li, L.  
TITLE Regulation of synaptophysin degradation by mammalian homologues of  
seven in absentia  
JOURNAL J. Biol. Chem. 277 (12), 10273-10282 (2002)  
MEDLINE 21895831  
PUBMED 11786535

REFERENCE  
AUTHORS Wheeler, T.C., Roudabush, F.L., Chin, L.-S. and Li, L.  
TITLE Direct Submission  
JOURNAL Submitted (08-JUN-2001) Department of Pharmacology, University of  
North Carolina at Chapel Hill, CB 7365, Chapel Hill, NC 27599, USA  
FEATURES  
source Location/Qualifiers  
1. .1465

/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strains="Sprague-Dawley"  
/db\_xref="taxon:10116"  
1. .849  
/codon\_start=1  
/product="SIAH-1A"  
/protein\_id="AAL91362.1"  
/db\_xref="GI:19550383"  
/translation="MSRQTALPTGTSKPPSQRPALPTGTSANNNDLASLFCPCV  
FDYVLPPILOQSGHLVCSNCRPKLTCCPTCRGLSGIRNLAMEKANSVLFPCKYAS  
SGCEITLPHTEKAEHEELCEFRPYSCPGASCKWQSGSLDAMPHMVRQHKSIITLQ  
EDIVFLATDINLPGADVVRMQSCFGEFHMVLVQKEXYGGHQFFAFVQLIGTRKQ  
ENFAIVELNGHRRRLTWEATPRSIHEGIATAIMNSDCLVFTSIAQLFAENGNLGIN  
VTISM"

ORIGIN  
Query Match 4.2%; Score 53; DB 10; Length 1465;  
Best Local Similarity 100.0%; Pred. No. 4e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1120 TTTGCAGAAATGGCAATTAGGCATCAATGTAACATTTCCATGTGTTGAAA 1172  
|||||  
Db 799 TTTGCAGAAATGGCAATTAGGCATCAATGTAACATTTCCATGTGTTGAAA 851  
|||||

RESULT 40  
AC105462 220157 bp DNA linear HTG 10-MAY-2003  
LOCUS Rattus norvegicus clone CH230-145D2, WORKING DRAFT SEQUENCE.  
DEFINITION AC105462  
ACCESSION AC105462  
VERSION AC105462.7 GI:30521365  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS 1 (bases 1 to 220157)  
Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebregeorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,  
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorensuhewa, L., Loulseghe, H., Lozard, R.J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwaokemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,  
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Welles, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 220157)  
Worley, K.C.

REFERENCE JOURNAL  
TITLE  
AUTHORS  
JOURNAL

Submitted (09-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 220157)  
Rat Genome Sequencing Consortium.

REFERENCE JOURNAL  
TITLE  
AUTHORS  
JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 10, 2003 this sequence version replaced gi:25008590.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GHPM  
Center clone name: CH230-145D2  
----- Summary Statistics  
Assembly program: Atlas 3.0:  
Consensus quality: 215295 bases at least Q40  
Consensus quality: 216317 bases at least Q30  
Consensus quality: 217167 bases at least Q20  
Estimated insert size: 223095; sum-of-contigs estimation  
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* By the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* the accession number will be 220157 bp in length.  
1 220157: contig of 220157 bp in length.  
Location/Qualifiers  
1. .220157  
/organism="Rattus norvegicus"

-----  
\* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* By the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* the accession number will be 220157 bp in length.  
1 220157: contig of 220157 bp in length.  
Location/Qualifiers  
1. .220157  
/organism="Rattus norvegicus"

/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-145D2"  
1. .997  
misc\_feature  
clone end: T7  
/note="wgs end extension"  
1295-2037  
misc\_feature  
clone end: T7  
/note="clone boundary"  
clone end: T7  
site: EcoRI  
end sequence: BH297589"  
ORIGIN  
Query Match 4.2%; Score 53; DB 2; Length 220157;  
Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1120 TTTCAGAAATGGCAATTTAGGCATCAATGTAACTATTTCATGTTGAAA 1172  
|||||  
Db 118973 TTTCAGAAATGGCAATTTAGGCATCAATGTAACTATTTCATGTTGAAA 119025  
AC098991 232382 bp DNA linear HTG 10-MAY-2003  
Rattus norvegicus clone CH230-84N8, WORKING DRAFT SEQUENCE, 2  
unordered pieces.  
AC098991  
AC098991.7 GI:30520816  
HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
Rattus norvegicus  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 232382)  
Muzny, D., Marie, Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Evans, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhen, L., Louleghed, H., Lozano, R., Lu, X., Ma, J., Mangum, B., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,

Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 232382)  
Worley, K.C.  
Direct Submission  
Submitted (08-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 232382)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 10, 2003 this sequence version replaced gi:25090763.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GHZA  
Center clone name: CH230-84N8  
----- Summary Statistics  
Assembly program: Atlas 3.0;  
Consensus quality: 222927 bases at least Q40  
Consensus quality: 225089 bases at least Q30  
Consensus quality: 226508 bases at least Q20  
Estimated insert size: 232565; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 230782: contig of 230782 bp in length  
\* 230783 230882: gap of unknown length  
\* 230883 232382: contig of 1500 bp in length.  
\* Location/Qualifiers  
1..232382  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL  
TITLE  
JOURNAL  
COMMENT

Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 232382)  
Worley, K.C.  
Direct Submission  
Submitted (08-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 232382)  
Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On May 10, 2003 this sequence version replaced gi:25090763.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GHZA  
Center clone name: CH230-84N8  
----- Summary Statistics  
Assembly program: Atlas 3.0;  
Consensus quality: 222927 bases at least Q40  
Consensus quality: 225089 bases at least Q30  
Consensus quality: 226508 bases at least Q20  
Estimated insert size: 232565; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 230782: contig of 230782 bp in length  
\* 230783 230882: gap of unknown length  
\* 230883 232382: contig of 1500 bp in length.  
\* Location/Qualifiers  
1..232382  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL  
TITLE  
JOURNAL  
COMMENT

misc\_feature  
1..1936  
/clone="CH230-84N8"  
/note="wgs\_contig"  
34268..35347  
/note="wgs\_contig"  
misc\_feature  
1..1936  
/clone="CH230-84N8"  
/note="wgs\_contig"  
34268..35347  
/note="wgs\_contig"  
ORIGIN  
Query Match  
Best Local Similarity 4.2%; Score 53; DB 2; Length 232382;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1120 TTTCAGAAAATGGCAATTTAGGCATCAATGTAACATTTTCATGTTGAAA 1172  
Db 5705 TTTCAGAAAATGGCAATTTAGGCATCAATGTAACATTTTCATGTTGAAA 5653  
RESULT 42  
CQ486802/c  
LOCUS CQ486802 426 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 18669 from Patent WO0160860.  
ACCESSION CQ486802  
VERSION CQ486802.1 GI:41452421  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 Schlegel, R., Endege, W.O. and Monahan, J.E.  
AUTHORS Genes differentially expressed in human prostate cancer and their  
TITLE use  
JOURNAL Patent: WO 0160860-A 18669 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)  
FEATURES  
source  
Location/Qualifiers  
1..426  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match  
Best Local Similarity 4.1%; Score 52; DB 6; Length 426;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 965 GCAAGCAAGCTGAAAATTTGCTTACCACCTTGAGCTAAATGTCATAGGCG 1016  
Db 426 GCAAGCAAGCTGAAAATTTGCTTACCACCTTGAGCTAAATGTCATAGGCG 375  
RESULT 43  
MMSIAHPSA  
LOCUS MMSIAHPSA 2123 bp DNA linear ROD 01-MAY-1993  
DEFINITION M.musculus pseudogene.  
ACCESSION Z19582  
VERSION Z19582.1 GI:297038  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 2123)  
AUTHORS Della.N.G., Senior, P.V. and Bowtell, D.D.  
TITLE Isolation and characterisation of murine homologues of the  
JOURNAL Drosophila seven in absentia gene (sina)  
REFERENCE Development (1993) In press  
2 (bases 1 to 2123)  
AUTHORS Bowtell, D.D.  
TITLE Direct Submission  
JOURNAL Submitted (15-JAN-1993) David D.L. Bowtell, Developmental Biology  
Group, Howard Florey, Institute of Exp. Physiology and Medicine,  
University of Melbourne, Parkville, Victoria, 3052, Australia  
FEATURES  
source  
Location/Qualifiers  
1..2123

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="Baib/c"
/db_xref="taxon:10090"
324..1166
/pseudo
/codon_start=1

ORIGIN
Query Match      4.1% Score 52; DB 10; Length 2123;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 395 TGACTGGCACTGATCCACCAATGACTTGGCGAGTCTTTTGGATGTC 446
      |||||||
Db 397 TGACTGGCACTGATCCACCAATGACTTGGCGAGTCTTTTGGATGTC 448

RESULT 44
AC139553/c
LOCUS
DEFINITION Mus musculus chromosome 19, clone RP24-386E18, complete sequence.
ACCESSION AC139553
VERSION AC139553.5 GI:45917411
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
TITLE Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
JOURNAL Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
AUTHORS Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
McDonald,J., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbbs,M.,
Talamas,J., Teefaye,S., Theodor,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (05-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 168091)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
McDonald,J., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbbs,M.,
Talamas,J., Teefaye,S., Theodor,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (14-JAN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 168091)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
McDonald,J., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbbs,M.,
Talamas,J., Teefaye,S., Theodor,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-APR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 1, 2004 this sequence version replaced gi:40841749.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
----- Project Information
Center project name: L21117
Center clone name: 386_E_18
-----
Some of the sequence contained within base pairs 1 - 85216 was
stolen from accession AC123941.
Location/Qualifiers
1..168091
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="19"
/map="19"
/clone="RP24-386E18"
/clone_lib="RPCI-24 Male Mouse BAC"
1..5217
/note="wgs end extension
clone end:SP6"
complement(785..909)
/rpt_family="B1P"
986..1164
/rpt_family="MER59A"
1554..2073
/rpt_family="RMER1B"
2074..2127
/rpt_family="(TG)n"
2128..2243
/rpt_family="RMER1B"
complement(3146..3420)
/rpt_family="B4A"
FEATURES
source
misc_feature
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

```

```
repeat_region 4694..4778
misc_feature /rpt_family="(GAAAA)n"
5218..5221
/note="clone boundary
clone end:SP6
site:Nboi"
repeat_region 5666..5941
/rpt_family="MERVL LTR"
repeat_region complement(5957..6239)
/rpt_family="MERVL"
repeat_region 6238..6970
/rpt_family="MERVL LTR"
repeat_region 6971..7159
/rpt_family="MERVL LTR"
repeat_region complement(7575..7698)
/rpt_family="MLTIF2"
repeat_region 8181..8213
/rpt_family="(TG)n"
repeat_region complement(8225..8332)
/rpt_family="RSINE1"
repeat_region complement(8369..8483)
/rpt_family="MIR"
repeat_region 8784..8955
/rpt_family="B3"
repeat_region 9041..9114
/rpt_family="A-rich"
repeat_region 9316..9451
/rpt_family="B1_MM"
repeat_region 9453..9483
/rpt_family="(CAAAA)n"
repeat_region complement(10015..10576)
/rpt_family="MER31-int"
repeat_region 11076..11248
/rpt_family="MER5C"
repeat_region 12991..13013
/rpt_family="(TTTG)n"
repeat_region 13169..13304
/rpt_family="GA-rich"
repeat_region 13371..13393
/rpt_family="(TTA)n"
repeat_region 13948..13977
/rpt_family="(TA)n"
repeat_region 13985..14012
/rpt_family="AT-rich"
repeat_region 14647..14683
/rpt_family="AT-rich"
repeat_region complement(15104..15421)
/rpt_family="ORR1D"
repeat_region 15434..15599
/rpt_family="Lx2"
repeat_region 17365..17437
/rpt_family="(CATA)n"
repeat_region 17532..17815
/rpt_family="Lx"
repeat_region 18009..18120
/rpt_family="B3"
repeat_region complement(18524..19070)
/rpt_family="Lx9"
repeat_region 19069..19459
/rpt_family="Lx9"
repeat_region 19760..19808
/rpt_family="(CA)n"
repeat_region complement(20381..20565)
/rpt_family="B3"
repeat_region 22263..22489
/rpt_family="L2"
repeat_region complement(23375..23649)
/rpt_family="MER58B"
repeat_region 24147..24195
/rpt_family="(TAGA)n"
repeat_region 25258..25431
/rpt_family="MIR"
repeat_region 25881..25903
```

```
Query Match 4.1%; Score 52; DB 10; Length 168091;
Best Local Similarity 100.0%; Pred.No. 8.3e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 395 TGACTGGCACAACTGCATCCACCAATGACTGGCGAGTCTTTTTCAGTGTCC 446
Db 48075 TGACTGGCACAACTGCATCCACCAATGACTGGCGAGTCTTTTTCAGTGTCC 48024

/rpt_family="(A)n"

RESULT 45
AC123941 182733 bp DNA linear ROD 01-JAN-2004
Mus musculus BAC clone RP23-458G24 from chromosome 19, complete
sequence.
AC123941
AC123941.3 GI:38229437
HTG.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Goyea.R., Cotton.M., Spalding.L., Haakenson.W. and Schatzkamer,K.
1 (bases 1 to 182733)
The sequence of Mus musculus BAC clone RP23-458G24
Unpublished (2001)
2 (bases 1 to 182733)
Wilson,R.
Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 182733)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (05-JUN-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 182733)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (20-SEP-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 182733)
Wilson,R.K.
Direct Submission
Submitted (08-NOV-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 182733)
Wilson,R.
Direct Submission
Submitted (01-JAN-2004) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Nov 8, 2003 this sequence version replaced gi:23238069.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
-----
Center project name: M_BA0458G24

COMMENT
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.
```

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

## SOURCE INFORMATION:

The RPCI-23 BAC library has been constructed by Kazutoyo Osegawa and Minako Tateno in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

## NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC108421

## FEATURES

## Location/Qualifiers

```
1. .182733
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /chromosome="19"
   /map="19"
   /clone="RP23-458G24"
   /clone_lib="RPCI-23"
   352. .452
   /rpt_family="Alu"
   repeat_region
   1255. .1439
   /rpt_family="MaLR"
   repeat_region
   1668. .1720
   /rpt_family="B2"
   repeat_region
   4179. .4767
   /rpt_family="RMER19B"
   repeat_region
   9359. .9608
   /rpt_family="B4"
   repeat_region
   9991. .10078
   /rpt_family="L2"
   repeat_region
   12166. .12213
   /rpt_family="ERV1"
   repeat_region
   12233. .12378
   /rpt_family="B2"
   repeat_region
   12422. .12851
   /rpt_family="RMER6A"
   repeat_region
   12954. .13423
   /rpt_family="RMER6A"
   repeat_region
   13683. .13787
   /rpt_family="B4"
   repeat_region
   14752. .14960
   /rpt_family="RMER6A"
   repeat_region
   15036. .15181
   /rpt_family="B4"
   repeat_region
   15226. .15350
   /rpt_family="Alu"
   repeat_region
   15687. .15840
   /rpt_family="L1"
   repeat_region
   15907. .16126
   /rpt_family="L1"
   repeat_region
   16210. .16401
   /rpt_family="L1"
   repeat_region
   16557. .16822
   /rpt_family="L1"
   repeat_region
   16824. .17207
   /rpt_family="L1"
   repeat_region
   17208. .17312
   /rpt_family="L1"
   repeat_region
   17311. .18615
   /rpt_family="L1"
   repeat_region
   18617. .19938
   /rpt_family="L1"
   repeat_region
   20051. .20917
   /rpt_family="L1"
```

```
repeat_region 21174. .21589
/rpt_family="L1"
repeat_region 21749. .22073
/rpt_family="MaLR"
repeat_region 22082. .22154
/rpt_family="ID"
repeat_region 22156. .22290
/rpt_family="B4"
repeat_region 22302. .22602
/rpt_family="L1"
repeat_region 22754. .23398
/rpt_family="L1"
repeat_region 23443. .23830
/rpt_family="L1"
repeat_region 26441. .26609
/rpt_family="B2"
repeat_region 26627. .27426
/rpt_family="L1"
repeat_region 28296. .28500
/rpt_family="B2"
repeat_region 28688. .28941
/rpt_family="L1"
repeat_region 29068. .29286
/rpt_family="L1"
repeat_region 30687. .31216
/rpt_family="L1"
repeat_region 31425. .32157
/rpt_family="L1"
repeat_region 33002. .33311
/rpt_family="L2"
repeat_region 34081. .34405
/rpt_family="ERVK"
repeat_region 36448. .36645
/rpt_family="L1"
repeat_region 36822. .36969
/rpt_family="Alu"
repeat_region 39590. .39805
/rpt_family="ERVK"
repeat_region 40632. .40774
/rpt_family="MIR"
repeat_region 40941. .41286
/rpt_family="MaLR"
repeat_region 42085. .42298
/rpt_family="B4"
repeat_region 44879. .45003
/rpt_family="L1"
repeat_region 45994. .46181
/rpt_family="B2"
repeat_region 46506. .46845
/rpt_family="MaLR"
repeat_region 46846. .47124
/rpt_family="MaLR"
repeat_region 48070. .48221
/rpt_family="L1"
repeat_region 48285. .48416
/rpt_family="B4"
repeat_region 52362. .53090
/rpt_family="L1"
repeat_region 53089. .53492
/rpt_family="L1"
repeat_region 53884. .54018
/rpt_family="L1"
repeat_region 54019. .54440
/rpt_family="ERVK"
repeat_region 54441. .55046
/rpt_family="L1"
repeat_region 55042. .55773
```

Query Match 4.1%; Score 52; DB 10; Length 182733;

Best Local Similarity 100.0%; Pred. No. 8.2e-16;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

395 TGACTGGCACAACTGCTCCATCCACAATGACTTGGCGAGTCTTTTGGTGTC 446

Db	37142	TGACTGCACAACTGCATCCAAATGACTTGCAGTCTTTTGTAGTGTC	37193	
RESULT 46				
LOCUS	MMSIAH1B	1713 bp	linear	MOD 03-NOV-1993
DEFINITION	M.musculus siah-1B protein mRNA.			
ACCESSION	Z19580			
VERSION	219580.1	GI:297801		
KEYWORDS	siah-1B protein.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE				
1.	(bases 1 to 1713)			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	Della, N.G., Senior, P.V. and Bowtell, D.D.			
JOURNAL	Isolation and characterisation of murine homologues of the			
MEDLINE	Drosophila seven in absentia gene (sina)			
PUBMED	94008536			
REFERENCE	2	(bases 1 to 1713)		
AUTHORS	Bowtell, D.D.			
TITLE	Direct Submission			
JOURNAL	Submitted (15-JAN-1993) David D.L. Bowtell, Developmental Biology			
FEATURES	source			
1.	.1713			
2.	/organism="Mus musculus"			
3.	/mol_type="mRNA"			
4.	/strain="random bred Swiss"			
5.	/db_xref="taxon:10090"			
6.	1_333			
7.	188*189			
8.	334..1182			
9.	/codon_start=1			
10.	/product="siah-1B protein"			
11.	/protein_id="CAA79631.1"			
12.	/db_xref="GI:297802"			
13.	/db_xref="GOA:Q06985"			
14.	/db_xref="UniProt/TREMBL:Q06985"			
15.	/translation="MSRQATALSTGSKPPSRQVPALDTTASNNDLASFCEPVC			
16.	FYVLPPIQCSGHVLSNCPKLTCTCPCTCRPLGSLRNLAEMKXVANSVLPCKYSA			
17.	SGCEITLPTKKAHEELCEPRYPSPCPGCKWQSLDAVMPHLMHOKHSITTLQ			
18.	EDIVFLATDINLPGAVDWMQSCFGEHMLVLEKQYDGHQOFPAIVQLIGTKQA			
19.	ENFAYRLENGHRRLLTWEATPSRHSIEGIATFAMNSDCLVFTSIQFLPAENGLGIN			
20.	VTISMV"			
21.	1183..1713			
22.	1692..1697			
23.	1712*1713			
3' UTR				
polyA_signal				
polyA_site				
ORIGIN				
Query Match	3.5%;	Score 45;	DB 10;	Length 1713;
Best Local Similarity	100.0%;	Pred. No. 6.7e-12;		
Matches 45;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	402	CACAACTGCATCCAAATGACTTGCAGTCTTTTGTAGTGTC	446	
Db	414	CACAACTGCATCCAAATGACTTGCAGTCTTTTGTAGTGTC	458	
RESULT 47				
LOCUS	BC052887	1720 bp	linear	MOD 30-JUN-2004
DEFINITION	Mus musculus seven in absentia 1B, mRNA (cdna clone MGC:60693			
ACCESSION	BC052887			
VERSION	BC052887.1	GI:31127269		
KEYWORDS	MGC.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE				
1.	(bases 1 to 1720)			
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,			
TITLE	Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,			
JOURNAL	Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,			
MEDLINE	Hopkins, R.P., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,			
PUBMED	Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,			
REFERENCE	Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,			
AUTHORS	Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,			
TITLE	Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,			
JOURNAL	Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,			
MEDLINE	McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,			
PUBMED	Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
REFERENCE	Fahay, J., Helton, E., Kettner, M., Madan, A., Rodrigues, S.,			
AUTHORS	Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
TITLE	Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,			
JOURNAL	Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,			
MEDLINE	Butterfield, Y.S., Kzywinski, M.I., Skalska, U., Smalil, D.E.,			
PUBMED	Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.			
REFERENCE	Generation and initial analysis of more than 15,000 full-length			
AUTHORS	human and mouse cDNA sequences			
TITLE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
JOURNAL	12477932			
2	(bases 1 to 1720)			
AUTHORS	Strausberg, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (27-MAY-2003) National Institutes of Health, Mammalian			
FEATURES	Gene Collection (MGC), Cancer Genomics Office, National Cancer			
1.	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
2.	USA			
3.	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>			
4.	Contact: MGC help desk			
5.	Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>			
6.	Tissue Procurement: Minoru Ko			
7.	CDNA Library Preparation: Yulan Piao and Minoru Ko (National			
8.	Institute on Aging, NIH: <a href="http://lgsun.grc.nia.nih.gov/cDNA/">http://lgsun.grc.nia.nih.gov/cDNA/</a> )			
9.	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
10.	DNA Sequencing by: National Institutes of Health Intramural			
11.	Sequencing Center (NISC),			
12.	Gaithersburg, Maryland;			
13.	Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a>			
14.	Contact: <a href="mailto:nisc.mgc@nih.gov">nisc.mgc@nih.gov</a>			
15.	Akter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,			
16.	Blakesley, R.W., Bouffard, G., Breen, K., Brinkley, C., Brooks, S.,			
17.	Districh, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,			
18.	Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,			
19.	Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,			
20.	McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,			
21.	Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,			
22.	Young, A., Zhang, L.-H. and Green, E.D.			
Clone distribution:	MGC clone distribution information can be found			
through the I.M.A.G.E. Consortium/LLNL at:	<a href="http://image.llnl.gov">http://image.llnl.gov</a>			
Series:	IRAK Plate: 112 Row: h Column: 15			
This clone was selected for full length sequencing because it				
passed the following selection criteria:	matched mRNA gi: 6677948.			
Location/Qualifiers				
1.	.1720			
2.	/organism="Mus musculus"			
3.	/mol_type="mRNA"			
4.	/strain="C57BL/6J"			
5.	/db_xref="taxon:10090"			
6.	/clone="MGC:60693 IMAGE:30053109"			
7.	/tissue_type="Egg, unfertilized, mouse"			
8.	/clone_lib="NIA Mouse Unfertilized Egg cDNA Library			
9.	(Long)"			
10.	/lab_host="DH10B"			
11.	/note="Vector: pSPORT1"			
12.	.1720			
13.	/gene="siah1b"			
14.	/db_xref="LocusID:20438"			



CDS

/db\_xref="MGI:108063"  
329..1177  
/gene="Shahib"  
/codon\_start=1  
/product="seven in absentia 1B"  
/protein\_id="AAH52887.1"  
/db\_xref="GI:31127270"  
/db\_xref="LocusID:20438"  
/db\_xref="MGI:108063"  
/translation="MSRQATALSTGTSKPPSRQVPALTDTTANNDLASLFCPVC  
FDYVLPPIQCSGHVSCNRPKLTCCPTCRGPIGSIRNLAVEKVANSVLPCKYSA  
SCGETLPHTKAEHELCEFPYPSCPCGASCKWGSGDGVMPHLMHOKHSITTLQ  
EDIVLATDINLPQAVDWMOQSCGFHMLVLEKQEKYDQHQFPAIVQLIGTRKQA  
ENFAYRLELNGHRRRLTWEATPRSHIEGATATMNSDCLVFTSIAQLPAENGNLGIN  
VTISM"

ORIGIN

Query Match 3.5%; Score 45; DB 10; Length 1720;  
Best Local Similarity 100.0%; Pred. No. 6.7e-12;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 402 CACAACGTGCATCAACAATGACTTGGCGAGTCTTTTGGAGTGTC 446  
|||||  
Db 409 CACAACGTGCATCAACAATGACTTGGCGAGTCTTTTGGAGTGTC 453  
|||||

RESULT 48  
AL732294/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AL732294 Mouse DNA sequence from clone RP23-59N7 on chromosome X, complete sequence.  
AL732294  
AL732294.18 GI:38568122  
HTG.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 112893)  
Clark, S.  
Direct Submission  
Submitted (27-NOV-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
On Nov 30, 2003 this sequence version replaced gi:38304169.  
Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
-----  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; SW.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at

/db\_xref="MGI:108063"  
329..1177  
/gene="Shahib"  
/codon\_start=1  
/product="seven in absentia 1B"  
/protein\_id="AAH52887.1"  
/db\_xref="GI:31127270"  
/db\_xref="LocusID:20438"  
/db\_xref="MGI:108063"  
/translation="MSRQATALSTGTSKPPSRQVPALTDTTANNDLASLFCPVC  
FDYVLPPIQCSGHVSCNRPKLTCCPTCRGPIGSIRNLAVEKVANSVLPCKYSA  
SCGETLPHTKAEHELCEFPYPSCPCGASCKWGSGDGVMPHLMHOKHSITTLQ  
EDIVLATDINLPQAVDWMOQSCGFHMLVLEKQEKYDQHQFPAIVQLIGTRKQA  
ENFAYRLELNGHRRRLTWEATPRSHIEGATATMNSDCLVFTSIAQLPAENGNLGIN  
VTISM"

ORIGIN

Query Match 3.5%; Score 45; DB 10; Length 1720;  
Best Local Similarity 100.0%; Pred. No. 6.7e-12;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 402 CACAACGTGCATCAACAATGACTTGGCGAGTCTTTTGGAGTGTC 446  
|||||  
Db 409 CACAACGTGCATCAACAATGACTTGGCGAGTCTTTTGGAGTGTC 453  
|||||

RESULT 48  
AL732294/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AL732294 Mouse DNA sequence from clone RP23-59N7 on chromosome X, complete sequence.  
AL732294  
AL732294.18 GI:38568122  
HTG.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 112893)  
Clark, S.  
Direct Submission  
Submitted (27-NOV-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
On Nov 30, 2003 this sequence version replaced gi:38304169.  
Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
-----  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; SW.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-59N7 is from the RPCI-23 Mouse BAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACE3.6.

FEATURES  
source  
1..112893  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="X"  
/clone="RP23-59N7"  
/clone\_lib="RPCI-23"

## ORIGIN

Query Match 3.5%; Score 45; DB 10; Length 112893;  
Best Local Similarity 100.0%; Pred. No. 4.3e-12;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 402 CACAACGTGCATCAACAATGACTTGGCGAGTCTTTTGGAGTGTC 446  
|||||  
Db 63845 CACAACGTGCATCAACAATGACTTGGCGAGTCTTTTGGAGTGTC 63801  
|||||

## RESULT 49

AC091606/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AC091606 234105 bp DNA linear ROD 21-MAY-2003  
Mus Musculus Strain C57BL6/J chromosome X BAC, RP23-75P20, complete sequence.  
AC091606  
AC091606.8 GI:30962723  
HTG.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 234105)  
Grills, G., Li, L., Montgomery, K.T., Chiu, D., Decker, J., Fusina, M., Goltz, J., Haider, A., Hall, L., Han, J., Ioshikhes, I.P., Lee, E., Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.  
High Throughput Mouse Sequencing  
Unpublished  
2 (bases 1 to 234105)  
Grills, G., Li, L., Montgomery, K.T., Chiu, D., Decker, J., Fusina, M., Goltz, J., Haider, A., Hall, L., Han, J., Ioshikhes, I.P., Lee, E., Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.  
Direct Submission  
Submitted (09-MAY-2001) Department of Molecular Genetics, Albert Einstein College of Medicine Genome Center, 1300 Morris Park Ave., Bronx, NY 10461, USA  
3 (bases 1 to 234105)  
Grills, G., Li, L., Montgomery, K.T., Chiu, D., Decker, J., Fusina, M., Goltz, J., Haider, A., Hall, L., Han, J., Ioshikhes, I.P., Lee, E., Perera, A., Shim, C., Thomas, E., Gendai, L.J. and Kucherlapati, R.  
Direct Submission  
Submitted (21-MAY-2003) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA  
On May 21, 2003 this sequence version replaced gi:18057068.  
-----Genome Center:  
Center: Harvard Partners Genome Center  
Center Code: HPGC  
Web site: <http://www.hpcgg.org/sequence/mouse.html>  
Contact: [hpgc@medel.mgh.harvard.edu](mailto:hpgc@medel.mgh.harvard.edu)

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550).  
Repeats are identified using RepeatMasker (A. Smit and P. Green,



```
repeat_region complement(18965..19338)
repeat_region /rpt_family="ORR1D"
repeat_region 19619..19682
repeat_region /rpt_family="PB1D9"
repeat_region complement(20370..20665)
repeat_region /rpt_family="MTC"
repeat_region 20665..20792

Query Match 3.5%; Score 45; DB 10; Length 234105;
Best Local Similarity 100.0%; Pred. NO. 4e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 402 CACACTGTCATCCACAATGACTTGGCGAGTCTTTTGAGTGTC 446
DB 57942 CACAACATGTCATCCACAATGACTTGGCGAGTCTTTTGAGTGTC 57898

RESULT 50
BV102520/c 523 bp DNA linear STS 28-FEB-2004
LOCUS MARC 14423-14424:1027951438:1 SCF - porcine spleen Sus scrofa STS
DEFINITION genomic, sequence tagged site.
ACCESSION BV102520
VERSION BV102520.1 GI:44680182
KEYWORDS STS.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 523)
AUTHORS Freking,B.A., Nonneman,D.P., Rohrer,G.A., Fahrenkrug,S.C.,
Smith,T.P.L. and Keeler,J.W.
TITLE Development of a comprehensive SNP-based comparative linkage map
for pigs
JOURNAL Unpublished (2004)
COMMENT

Contact: Freking BA
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4278
Fax: 402 762 4173
Email: freking@email.marc.usda.gov
Primer A: AIGCCTCTCTGGAGTGAGGA
Primer B: CGCCGCTATGACCATTTAG
STS size: 400
PCR Profile:
Hotstart: 95 degrees for 15 minutes
Denature: 95 degrees for 30 seconds
Anneal: 56 degrees
Extension: 68 degrees for 2 minutes
Cycles: 32 to 45
Protocol:
Template: 50-200 ng genomic DNA
Primer: each 20 pmoles
dNTPs: each 88 uM
Taq Polymerase: 0.25 units (Qiagen HotStar)

Buffer: Commercially supplied Qiagen HotStar buffer

The STS is derived from PCR amplicons generated from genomic DNA,
sequenced from each end using the amplification primers. The
sequence does not necessarily represent the entire amplicon.
Sequence derived from PolyPhred was trimmed from each end of each
unique contig until five consecutive bases exceeded a quality score
threshold of 20, and the next ten bases averaged a quality score of
20 or greater. Amplicon size was estimated by agarose gel
electrophoresis.

FEATURES
source 1..523
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="genomic DNA"
/strain="white composite, duroc, meishan, minzhu,
```

```
fengjing, crossbreeds"
/db_xref="taxon:9823"
/sex="male and female"
/clone_lib="SCF - porcine spleen"
/dev_stage="adult"
/note="Organ: spleen"
124..523
124..144
complement(504..523)

primer_bind
primer_bind
ORIGIN

Query Match 2.9%; Score 37; DB 11; Length 523;
Best Local Similarity 100.0%; Pred. NO. 1.3e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 TGAAGAGCTCTGTGAGTTTAGGCCTTATTCCTGTCG 708
DB 118 TGAAGAGCTCTGTGAGTTTAGGCCTTATTCCTGTCG 82

Search completed: April 25, 2005, 02:39:49
Job time : 5791 secs
```

**This Page Blank (uspto)**